5-1-2009

Color difference formula and uniform color space modeling and evaluation

Shizhe Shen

Color Difference Formula and Uniform Color Space Modeling and Evaluation

by

Shizhe Shen

B.S. Zhejiang University, Hangzhou, CHINA (2002)

M.S. Zhejiang University, Hangzhou, CHINA (2005)

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Color Science in the Center for Imaging Science, Rochester Institute of Technology

May, 2009

_____

Signature of the Author

_____

Accepted by Dr. Mark D. Fairchild,
Coordinator, M.S. Degree Program

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NY

CERTIFICATE OF APPROVAL

M.S. DEGREE THESIS

The M.S. Degree Thesis of Shizhe Shen
has been examined and approved by two members of the
Color Science faculty as satisfactory for the thesis
requirement for the Master of Science degree

Dr. Roy S. Berns, Thesis Advisor

Dr. Mark D. Fairchild

Color Difference Formula and Uniform Color Space Modeling and Evaluation

Shizhe Shen

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Color Science in the Center for Imaging Science, Rochester Institute of Technology

Abstract:

Defining color tolerances numerically continues to be a topic of intense interest in colorimetry. A technique was developed to evaluate formula performance that incorporated visual uncertainty. In this technique, visual uncertainty was represented by randomized equal color-difference ellipsoids or randomized visual color differences. STRESS, a multivariate statistical tool, was employed to quantify these randomized equal color-difference ellipsoids or visual color differences. The STRESS clouds were composed of the STRESS values between the randomized equal color-difference ellipsoids and T50 equal color-difference ellipsoids, or between the randomized visual color differences and T50 visual color differences where T50 represented visually determined tolerances equivalent to an anchor-pair stimulus. These STRESS values clouds were taken as rulers to evaluate whether one color-difference formula over-, under- or well-fitted a specified color-difference dataset, based on an F-test. This technique is a necessary addition to the current deviation evaluation metrics, e.g., PF/3. In follow-on research, a Euclidean color space was developed with the color-difference formula based on IPT color space for supra-threshold color differences. The color-difference formula has similar chromatic modeling to CIE94. A lightness transformation function was applied to model color difference along lightness. A rotation matrix on the chromatic plane was also applied to achieve better characteristics of the color space. A step-wise optimization was performed to achieve better consistency and remove conflicts between different color-difference datasets. The evaluations include STRESS, F-test, hue constancy and equal color-difference ellipsoid shape. It was shown by the evaluation results that the Euclidean color space could be a potential candidate of a future color model useful for defining industrial color tolerances.

i. DEDICATION

THIS THESIS IS DEDICATED TO MY FAMILY FOR THEIR LOVE AND SUPPORT.

ii. ACKNOWLEDGEMENTS

I would like to thank the following people for their help with and support for my thesis and my study in color science:

Roy S. Berns, my thesis advisor, for providing me the opportunity to study in Munsell Color Science Laboratory, finding my advantages and weaknesses, sharing his knowledge and vision on color science and guiding me exploring the amazing color world,

Mark D. Fairchild, for sharing his knowledge in color appearance models, extending the concept of color space in my brain from three-dimension to more dimensions,

Dave R. Wyble, for sharing his knowledge of photometry and spending time around ping-pong table with me,

Lawrence Taplin, for leading me into the wonderful Matlab world, helping me on all kinds of computer and programming questions and bringing old fashion pizza,

Mitchell R. Rosen, for introducing me the real world color management and advising me on several important periods,

Rod Heckaman, for his optimism and encouragement on me,

Val Hemink, for her hospitality and continuous help,

the Munsell Color Science Laboratory faculties, staffs and students, for teaching me many interesting things and giving me another warm family,

the DuPont Color Science Fellowship for their financial support of my research,

my friends and family, for their love and support,

and Ying, for her love, support and patience.

iii. TABLE OF CONTENTS

iv. LIST OF FIGURES

V. LIST OF TABLES

1. Introduction

There are four questions in color science: how to describe color, how to measure color, how to measure color quality and how to produce correct color [Berns, 2000]. In this thesis, all the efforts were applied to looking for answers for the third question: how to measure color quality. The solution is to build the relationship between numbers and visual color differences, that is, to model human perception of color [Green, 2002]. These models are so-called color-difference formulas or color-difference equations.

Before answering this question, one must first answer another question in the list: how to describe color in a numerical way. All the color-difference formulas must be based on certain color spaces (or color systems). In 1931, a color-matching system, tristimulus values, based on Guild and Wright's experiments [Guild, 1931] [Wright, 1928-1929], [Wright, 1929-1930], [Wright, 1930] was adopted by CIE and became the root of CIE colorimetry [CIE publ. 15, 2004]. From then on, continuous efforts were applied to look for better visually uniform color spaces and color-difference formulas [Hunter, 1987], [Berns, 2000]. After its debut in 1976, CIELAB soon became the most popular uniform color space [CIE publ. 15, 2004] and the Euclidean distance in CIELAB is widely used as the measure of color quality, even today. Most modern color-difference formulas were developed based on CIELAB, e.g., CMC [Clarke, 1984], BFD [Luo, 1987a], CIE94 [Berns, 1993] and CIEDE2000 [Luo, 2001], [CIE Publ. 15, 2004], after realizing the non-uniform nature of CIELAB revealed by a series of visual color-difference experiments. Among these visual color-difference experiments, four color-difference datasets were commonly used to model and evaluate small color-difference formulas: BFD-P [Luo, 1986], RIT-DuPont [Berns, 1991], Leeds [Kim, 1997] and Witt [Witt, 1999]. These

visual color-difference experiments were performed following the CIE guidelines on color-difference formula research [Robertson, 1978], [Witt, 1995]. The BFD-P and RIT-DuPont datasets are the most comprehensive datasets derived with spatially well-sampled color centers in color space. Furthermore, the BFD-P dataset was built on combining previous datasets with an anchor experiment. In 1998, Qiao's hue suprathreshold color-difference dataset was derived to extend the RIT-DuPont dataset emphasizing hue dependency [Qiao, 1998] and used in modeling the hue dependency function for CIEDE2000 [Berns, 2002]. The Qiao dataset could be combined with the RIT-DuPont dataset as a more comprehensive color-difference dataset. The latest CIE recommended color-difference formula, CIEDE2000, fit these small color-difference datasets very well but it has a quite complex form. Thus, people began to question the capability of CIELAB in representing small to medium color differences and tried to find color-difference formulas that have similar or better performance, but a simpler form than CIEDE2000 based on color spaces other than CIELAB [Li, 2002], [Cui, 2002], [Li, 2003], [Luo, 2006], [Huertas, 2006], [Berns, 2007], [Oleari, 2008], [Berns, 2008]. German standard DIN99 [DIN 6176, 2000], color appearance models CIECAM97s [CIE Publ. 131, 1998] and CIECAM02 [CIE Publ. 159, 2004], and OSA UCS [MacAdam, 1974] were considered as potential candidates and the DIN99d [Cui, 2002], DIN99o [Witt, 2009], CAM02 -SCD, -LCD and -UCS [Luo, 2006], CIECAM02 series [Berns, 2007], IPT series [Xue, 2008], ΔE_{GP} [Huertas, 2006], and ΔE_E [Oleari, 2008] were developed based on these since 2000. Also, new color spaces were generated for the purpose of color quality measure based on the extended understanding of human color vision [Berns, 2008]. These color-difference models have similar performance in statistical evaluation.

It is also necessary to discuss the relationship between color spaces, color appearance models and color-difference formulas. They were different research areas in the past, but became more and more interdisciplinary in recent color science research. Generally speaking, the Euclidean distance of a numerical color system can always be used as a color-difference formula, e.g., ΔE^*_{ab}. The reverse became true recently employing the technique to derive Euclidean color spaces from the clue provided by the color-difference formulas, either by analytical methods [Völz, 1998], [Völz, 1999-2000], [Thomsen, 2000], [Völz, 2006] or computational methods [Urban, 2007a], [Urban, 2007b]. CIELAB space was modified fully or partially based on different CIELAB based color-difference formulas, e.g., CMC [Völz, 1999-2000], CIE94 [Völz, 1998], [Thomsen, 2000] and CIE2000 [Völz, 2006], [Urban, 2007a], [Urban, 2007b]. The analytical method is to integrate the color-difference formula along certain colorimetric attributes, which let the Euclidean distance in new derived Euclidean space equal to the visual color difference represented by the color-difference formulas. While in the computational methods, look up tables were employed to map the relationship between the Euclidean color spaces and the color-difference formulas. In addition, the concept was adopted in directly modeling new uniform color spaces [Li, 2002], [Luo, 2006], [Berns, 2007], [Berns, 2008]. Even though the technique is still imperfect and not all the color-difference formulas can be exactly transformed to the corresponding Euclidean color spaces, it is obvious that color spaces and color-difference formulas began to merge together, functionally. On the other hand, the color appearance models, which were defined to model how the human visual system perceives the color of an object under different viewing conditions and with difference backgrounds [Fairchild, 2002], were

gradually employed as color spaces to describe color and measure color quality with their embedded three-dimensional colorimetric attributes [Luo, 1996], [Li, 2002], [Li, 2003], [Luo, 2006], [Berns, 2007]. The reverse is also true for most current color spaces that satisfied the definition by TC1-34: "…must account for at least chromatic adaptation and have correlates of at least lightness, chroma and hue." [Fairchild, 2002] The merge of color spaces, color-difference formulas and color appearance models predicts that there should be one universal color model that can predict, describe and measure perceptual color in "all" applications. Such a color model should at least have the following properties: (a) a visually uniform color space to describe color, and its Euclidean distance to measure color quality, defined in a reference environment (functions of color spaces and color-difference formulas); (b) a model to transform color appearance under different viewing conditions and with different background to the colorimetric attributes under the environment defined for this uniform color space (responsibility of color appearance model). Actually, the prototype of the model can be found in the current ICC color management pipeline, which adopts tristimulus values or CIELAB as the profile connection space (PCS) to describe color and matrices and/or look up tables to transform between different color spaces to the PCS [ICC specification, 2004].

The evaluation metrics are important and necessary for deriving better color models. Statistical functions calculating the deviation between visual color differences and numerical color differences calculated by color-difference formulas are the main numerical metrics to evaluate color-difference formulas. These functions include the correlation coefficient r, CV and γ proposed by Coates et al. [Coates, 1981] and V_{AB} derived by Schultz [Schultz, 1972]. These four functions were combined as one

performance factor (PF) in 1987 [Luo, 1987b]. In 1999, by realizing that a correlation coefficient *r* was quite inconsistent with the other measures, PF/3 was introduced and the correlation coefficient was removed from the list [Guan, 1999]. From then on, PF/3 became the most popular color-difference formula evaluation metric [Luo, 2001], [Melgosa, 2004]. Usually, the comparison between different color-difference formulas was also performed by a statistical test at the same time, e.g., F-test on V_{AB}. In 2007, a statistical index named STandardized REsidual Sum of Squares (STRESS), which is a multidimensional scaling metric, was recommended to evaluate the performance of different color-difference formulas [Garcia, 2007]. It combines features of Kruskal's Stress-1 and V_{AB}. Garcia's STRESS (hereafter called STRESS) has the advantage of making statistical inferences easily compared with other statistical metrics and has been successfully applied in current color-difference formula research [Melgosa, 2008], [Oleari, 2008].

It is another important supplemental evaluation metric that the comparison between the equal color-difference ellipsoids fitted with small color-difference datasets and the ellipsoids represented by the color-difference formulas [Luo, 2001], [Oleari, 2008]. For evaluating uniform color spaces, these equal color-difference ellipsoids are plotted in these color spaces and the uniformity of the shape and orientation of these ellipsoids is employed as the evaluation metric [Luo, 2002], [Luo, 2006]. The theoretical basis of this technique is that the equal color-difference contour of human vision can be represented by an ellipsoid, which is still a reliable assumption. Ellipses and ellipsoids have been widely used to represent equal visual color-difference contours [Alder, 1982], [Rich, 1983], [Cheung, 1986], [Luo, 1986], [Witt, 1990], [Melgosa, 1997] since the early

work of Judd [Judd, 1936], MacAdam [MacAdam, 1942], and Wright [Wright, 1943]. The fitting procedures for surface colors by the method of constant stimuli were well described in 1991 [Indow, 1991]. In 1997, the detailed procedures of fitting the RIT-DuPont dataset in CIELAB, derived based on probit analysis, were introduced and considered as the common technique [Melgosa, 1997]. Besides visual comparison, the numerical evaluation was introduced to compute the mismatch between ellipses [Rich, 1983], [Strocka, 1983]. The BFD-P and RIT-DuPont datasets are usually used for this purpose because the color space was spatially well sampled by these two datasets [Luo, 2001].

Hue constancy is another desired property of uniform color spaces. That is, in a color space with good hue constancy, the color attributes with the same hue angle should have the same perceptual hue. Some empirical color systems for large color differences have quite good hue constancy, e.g., Munsell color system [Wyszecki, 1982], OSA Uniform Color Scale (UCS) [MacAdam, 1974] and the Swedish Natural Color System (NCS) [Hård, 1996a], [Hård, 1996b]. But, the color matching based color spaces (e.g. CIELAB) and color appearance models (e.g. CIECAM97s and CIECAM02) do not maintain hue constancy very well. There were very few hue constancy datasets before the 1990's. The omission was filled by two constant hue visual experiments performed on CRT color in the 1990's [Hung, 1995] [Ebner, 1998a]. As a result, in 1998, IPT, a color space to preserve the hue constancy, was introduced, which predicts hue without detrimentally affecting other color appearance attributes [Ebner, 1998b].

In the second chapter of this thesis, several milestones in the history of uniform color space and color-difference formula modeling related to this research are briefly

introduced. The third chapter is composed of two research projects that are closely related to each other. They are presented in the form of three article drafts: in the first two parts, a new evaluating metric for color-difference formulas incorporating the visual uncertainty of the visual color-difference datasets was developed and introduced. This part of the research was triggered by realizing that the statistical function, e.g., PF/3 or STRESS, can only evaluate the prediction of the color-difference formulas to average visual data without considering the visual uncertainties inherent in the visual experiments. Even if the STRESS values of one color-difference formula for one color-difference dataset are small, it is possible the color-difference formula has fitted the error of the visual experiments. The perfect fit will be achieved only when the color-difference formula has equivalent performance to the uncertainty of the color-difference dataset [Witt, 1987]. In the third part, a Euclidean color space was derived from an optimized color-difference formula using the IPT color space with the analytical method. The IPT space was selected as the potential color space for developing a color-difference formula because of its simple form, easy implementation and good color appearance modeling. The Euclidean color space and color-difference formula were also evaluated and compared with others in the third part of chapter three.

2. Background

2.1. Small Color-Difference Datasets

A COM-weighted dataset was widely used in recent color-difference formula modeling and evaluating for small color difference [Luo, 1986] , [Berns, 1991], [Kim, 1997], [Witt, 1999]. The COM-weighted dataset was composed of four separate datasets developed in four different laboratories: the RIT-DuPont, BFD-P, Leeds and Witt datasets. Previously, these four datasets were used together and they were weighted based on the numbers of their color pairs to balance their importance in the combined dataset and they were adjusted to the same anchor pair [Luo, 2001]. But, it was found later that there are inner differences between these four datasets and they should be considered separately to avoid omitting important information [Berns, 1993], [Melgosa, 2004]. In this research, the RIT-DuPont dataset was combined with Qiao's hue suprathreshold color-difference dataset as another independent dataset [Qiao, 1998]. The summary of these datasets is shown in Table 2.1.

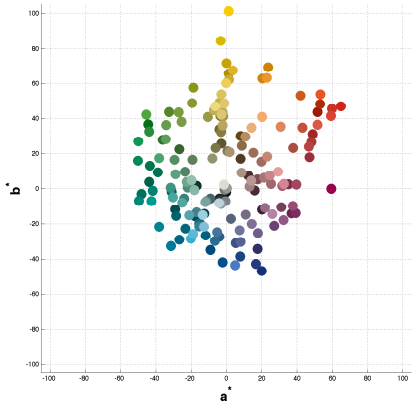
Table 2.1: Summary of the RIT-DuPont-Qiao, BFD-P, Leeds and Witt datasets.

Datasets	No. of pairs	Color difference range (ΔE_{ab}^*)	Pairs with $0.5 \le \Delta E_{ab}^* \le 5$			Sample materials	Psychophysical methods	Reference white point (XYZ)	Visual uncertainty
			% in all pairs	Range of lightness (CIELAB)	Range of chroma (CIELAB)				
RIT-DuPont-Qiao:	400*	0.78-4.41	100.00%	13.09 - 84.70	0.14 - 86.13	Glossy paint	Pair comparison	[94.81,100.00,107.30]	5.5%
RIT-DuPont	312*	0.78-4.41		13.09-84.70	0.14 - 86.13				
Qiao	88*	0.78-3.52		40, 60	20, 35	Glossy photo paper			11.5%
BFD-P:	2776	0.04-18.21	74.64%	0.90 - 92.76	0.06-104.62	Various materials and methods but relative scales of individual sets adjusted using textile samples and gray scale method.			8.9%
D65	2028	0.04-16.08	81.16%	0.90 - 92.76	0.06 - 85.23			[94.81,100.00,107.30]	
M	548	0.05-18.21	52.19%	26.88 - 77.27	0.23 - 92.14			[94.65,100.00,103.97]	
C	200	0.07-3.90	70.00%	15.22 - 87.68	3.53 - 104.62			[98.07,100.00,118.23]	
Leeds:	307	0.40-4.74	99.02%	27.35 - 90.22	0.19 - 50.25	Glossy paint	Gray scale and pair comparison	[94.81,100.00,107.30]	10%
Witt's:	418	0.12-10.63	86.60%	34.31 - 88.86	0.11 - 127.75	Glossy paint	Gray scale	[94.81,100.00,107.30]	10%

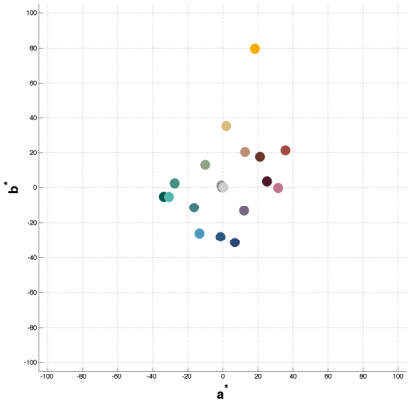
*: Both +T50 and –T50 are included.

In these experiments, different materials and psychophysical methods were employed, which indicates different parametric coefficients are necessary in color-difference modeling for an individual dataset. The RIT-DuPont and BFD-P datasets have the largest range of lightness and chroma. The research in this thesis focuses on small color-difference formula modeling; thus, only the color pairs with color difference within the range of $0.5 \leq \Delta E_{ab}^{*} \leq 5.0$ in the BFD-P, Leeds and Witt datasets were used in the modeling. The RIT-DuPont and Qiao experiments were designed for small color difference modeling, so all of the color vectors were within the small color-difference range recommended by the CIE [Robertson, 1978], [Witt, 1995].

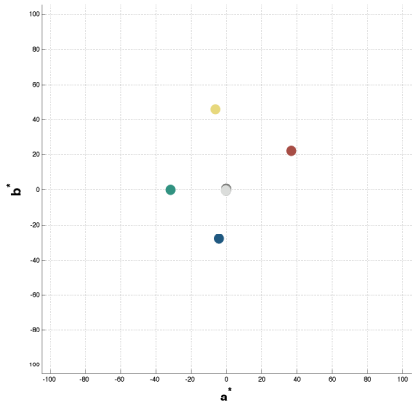
The color centers of the RIT-DuPont, BFD-P, Leeds and Witt datasets are illustrated in Figure 2.1. In Figure 2.1, the color centers of the BFD-M and BFD-C were transformed to the colorimetric attributes under D65 using a chromatic adaptation transformation. (Refer to Chapter 5 for details.) Also, hierarchical cluster analysis was performed on the Leeds, Witts and BFD-P datasets after chromatic adaptation to combine these color centers close enough to each other. (Refer to Chapter 4 for details.)



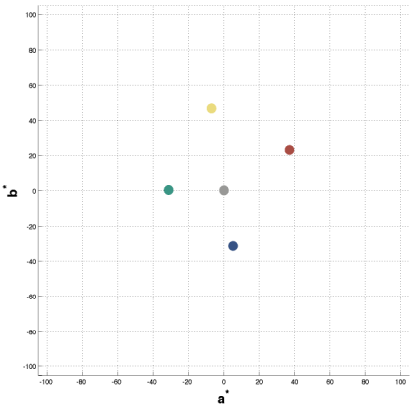
(a) BFD-P dataset



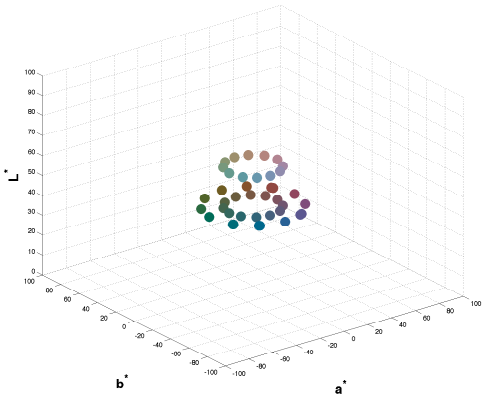
(b) RIT-DuPont dataset



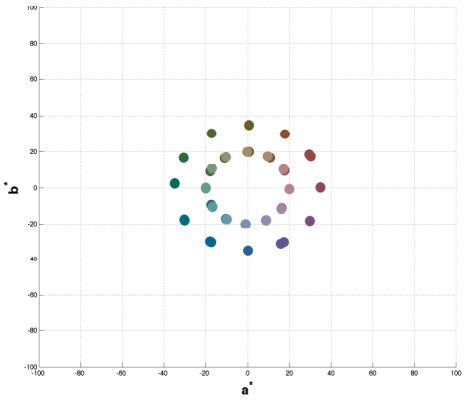
(c) Leeds dataset



(d) Witt dataset



(e) Qiao dataset (3-D view)



(f) Qiao dataset (a^*-b^* plane)

Figure 2.1: Color centers of small color-difference datasets. Color-coding is based on sRGB. [ICC, 1996]

2.2. Evaluation metrics for color-difference formulas and color spaces

2.2.1. PF/3

The original PF (Performance Factor) measuring the deviation between numerical color difference "ΔE" and the visual color difference ΔV was introduced in 1987 [Luo, 1987b], and shown in Equation (2.1). PF is the combination of four statistical measures in order to achieve a single value that represents different statistical evaluation.

$$PF = 100 \cdot (\gamma + V_{AB} + CV/100 - r) \tag{2.1}$$

where r represents the correlation coefficient, CV and γ were proposed by Coates et al. [Coates, 1981]. V_{AB} was derived by Schultz [Schultz, 1972]. They are given in Equations (2.2) – (2.5).

$$r = \frac{N\sum(X_i Y_i) - \sum X_i \sum Y_i}{\sqrt{\left[N\sum X_i^2 - \left(\sum X_i\right)^2\right]\left[N\sum Y_i^2 - \left(\sum Y_i\right)^2\right]}} \tag{2.2}$$

$$CV = \frac{\sqrt{\frac{1}{N}\sum\left(X_i - fY_i\right)^2}}{\overline{X}} \times 100 \tag{2.3}$$

with $$f = \frac{\sum X_i Y_i}{\sum Y_i^2}$$

$$\log_{10}(\gamma) = \sqrt{\frac{1}{N}\sum\left(\log_{10}\left(\frac{X_i}{Y_i}\right) - \log_{10}\left(\frac{X_i}{Y_i}\right)\right)^2} \tag{2.4}$$

$$V_{AB} = \sqrt{\frac{1}{N}\sum \frac{\left[X_i - (FY_i)\right]^2}{X_i F Y_i}} \tag{2.5}$$

with $$F = \sqrt{\sum \frac{X_i}{Y_i} \Big/ \sum \frac{Y_i}{X_i}}$$

where X_i, Y_i are two sets of data with same number N. In evaluating color-difference formulas, they are visual color differences and numerical color differences calculated by color-difference formulas for the same dataset.

In 1999, by realizing the correlation coefficient r was quite inconsistent with the other measures, PF/3 was introduced [Guan, 1999] and shown in Equation (2.6).

$$PF/3 = 100 \cdot \left[(\gamma - 1) + V_{AB} + CV/100\right] \Big/ 3 \tag{2.6}$$

For no errors, $CV = 0$, $V_{AB} = 0$, $\gamma = 1$ and PF/3 = 0. PF/3 was widely used as the measure of deviation in evaluating color-difference formulas [Luo, 2001], [Melgosa, 2004].

2.2.2. STRESS (STandardized REsidual Sum of Squares)

In 2007, a statistical index named standardized residual sum of squares (STRESS), which is based on Kruskal's Stress-1 multidimensional scaling metric, was recommended to evaluate the performance of color-difference formulas [Garcia, 2007]. The definition of STRESS for measuring the deviation between visual color differences and numerical color differences is:

$$STRESS = 100 \cdot \left(\frac{\sum_i \left(\Delta E_i - F \Delta V_i \right)^2}{\sum_i F^2 \Delta V_i^2} \right)^{1/2} \tag{2.7}$$

with $$F = \frac{\sum_i \Delta E_i^2}{\sum_i \Delta E_i \Delta V_i}$$

where ΔE defines the numerical color difference, ΔV defines the visual color difference, F is a scaling factor for adjusting ΔV values to the same scale as ΔE, and i is the number of color-difference pairs. Essentially, it is a normalized sum-of-squares error metric.

STRESS has the reported advantage of making statistical inferences compared with other metrics. For example, if two different sets of numerical color differences calculated by color-difference equations A and B have the same number of color pairs, an F-test between these two color-difference equations can be easily performed using each equation's STRESS values:

$$F_{value} = \frac{STRESS_A^2}{STRESS_B^2} \tag{2.8}$$

where $STRESS_A$ and $STRESS_B$ are the STRESS values for measuring the deviation between color-difference equations A and B and visual color differences. The hypothesis that color-difference equations A and B are not significantly different will be rejected if $F_{value} < F_C$ or $F_{value} > 1/F_C$, where F_C is the critical values of the F distribution with certain confidence level and (j-1, j-1) degrees of freedom, where j is the total number of color difference pairs. F_C values can be calculated using the Matlab function of *finv((1-0.95)/2, j-1, j-1)*, if the confidence level is 95%. STRESS was used in this research to

measure the deviations between visual color differences and numerical color differences computed by color-difference formulas, between two equal color-difference ellipsoids, and also between two sets of visual color differences. In addition, the comparison of color-difference formulas and color spaces was performed based on the F-test implemented with STRESS. One note of caution; differences are assumed to sample ratio scales, both having a true zero point.

2.2.3. Hue constant datasets

Both Hung and Berns' [Hung, 1995] and Ebner and Fairchild's [Ebner, 1998a] experiments were based on CRT colors. In Ebner and Fairchild's experiment, 306 points were sampled in CIELAB over 15 equally spaced hue angles (every 24 degrees), an average of 20 points in each reference hue plane. Thirty observers participated in the experiment. A total of 132 samples for 12 reference hues were evaluated by nine observers in Hung and Berns' experiment. For each reference hue, there were three samples at constant lightness and eight samples with varying lightness. In addition, in 2003, a radial sampling of data points in OSA-UCS were computed [Moroney, 2003]. These three hue constant datasets were used to evaluate the hue constancy for both color spaces and color appearance models [Hung, 1995], [Ebner, 1998a], [Moroney, 2003]. These hue constant datasets are plotted in CIELAB, CIECAM02 *Jab* based on *JCh*, IPT color spaces and OSA-UCS in Figures 2.2 – 2.5. Good hue constancy was achieved for IPT and OSA-UCS space, but not for CIELAB and CIECAM02 *Jab*.

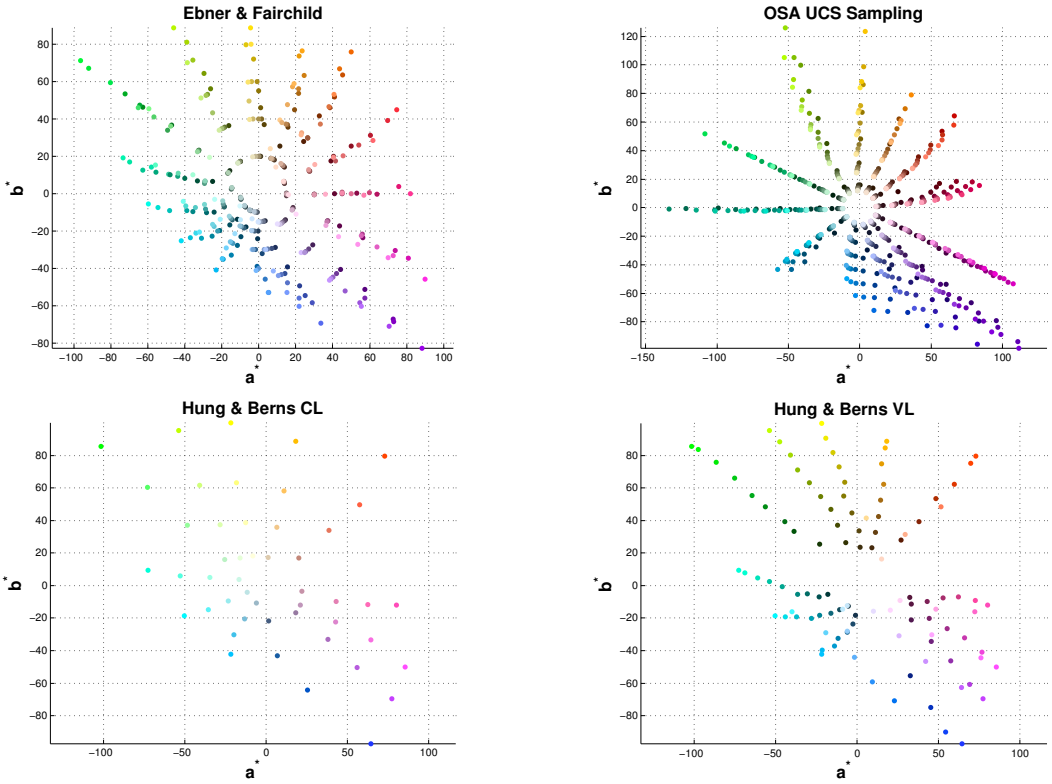


Figure 2.2: Constant hue datasets in CIELAB.

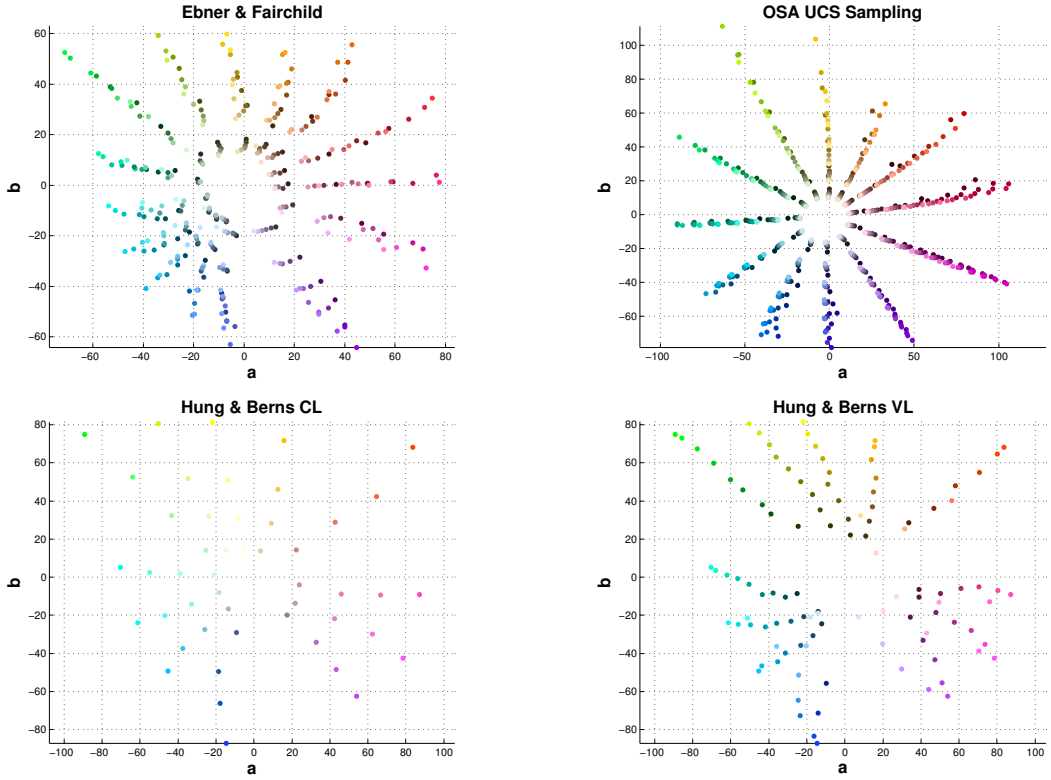


Figure 2.3: Constant hue datasets in CIECAM02.

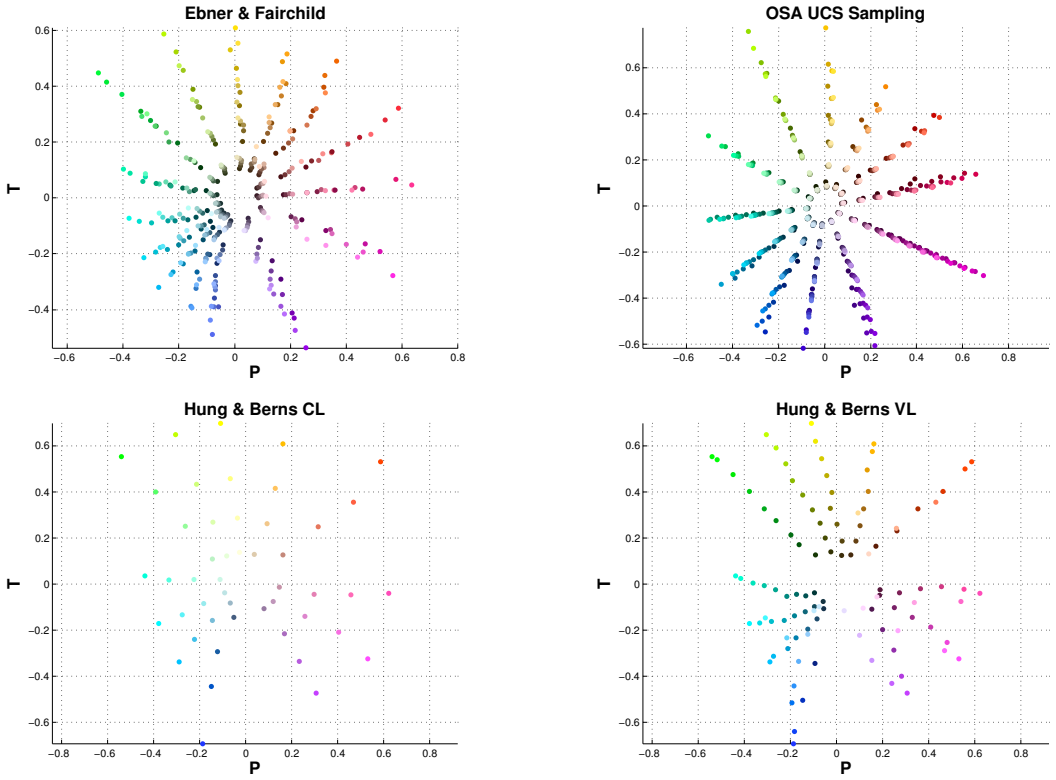


Figure 2.4: Constant hue datasets in IPT.

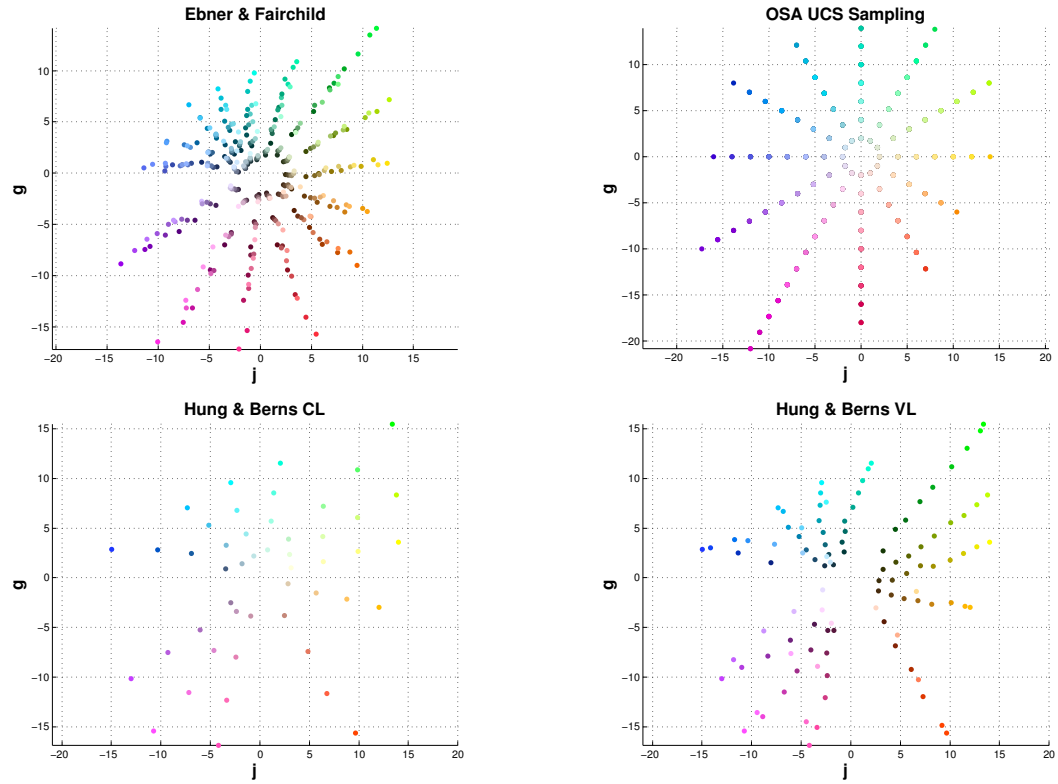


Figure 2.5: Constant hue datasets in OSA-UCS.

2.3. Equal Color-Difference Ellipsoids

The equal color-difference ellipsoids for the RIT-DuPont datasets were fitted and plotted on chromatic planes of CIELAB, CIECAM02 *Jab* based on *JCh* and IPT color spaces in Figures 2.6 – 2.8. The same fitting procedures used by Melgosa, et al. was followed [Melgosa, 1997]. For each color center, both the intersection (red curves) and projection (black curves) are plotted with three-times zooming. The mismatch between these two ellipses represents the direction of ellipsoids in lightness direction. In Figure 2.8, P and T are enlarged to achieve the colorimetric scale roughly close to CIELAB. It was shown in the Figure 2.6 there is a dependency of lengths along chroma increase on chroma and an obvious rotation happens in the blue area of CIELAB. The previous color-difference formulas were developed based on these findings [Berns, 1993], [Luo, 2001]. In Figures 2.7 and 2.8, the rotation of ellipses in blue area is improved and all the ellipses are pointed roughly towards the center at CIECAM02 and IPT spaces. The ellipses in IPT space still show the dependency on chroma but not for those in CIECAM02 space.

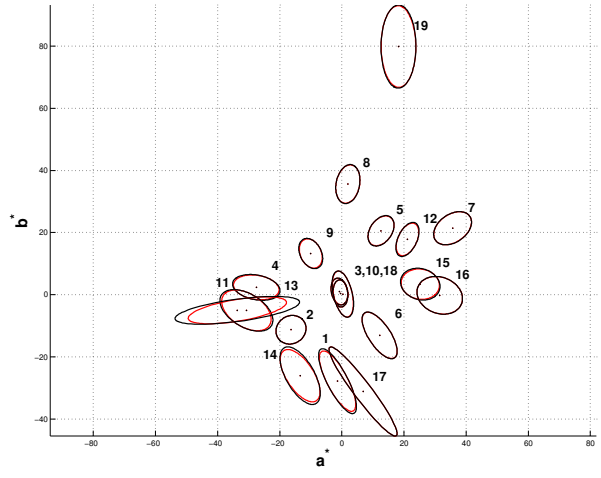


Figure 2.6: Equal color-difference ellipsoids of the RIT-DuPont datasets in CIELAB.

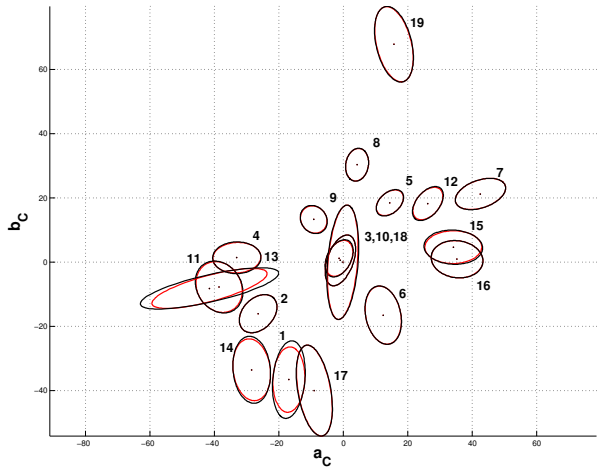


Figure 2.7: Equal color-difference ellipsoids of the RIT-DuPont datasets in CIECAM02.

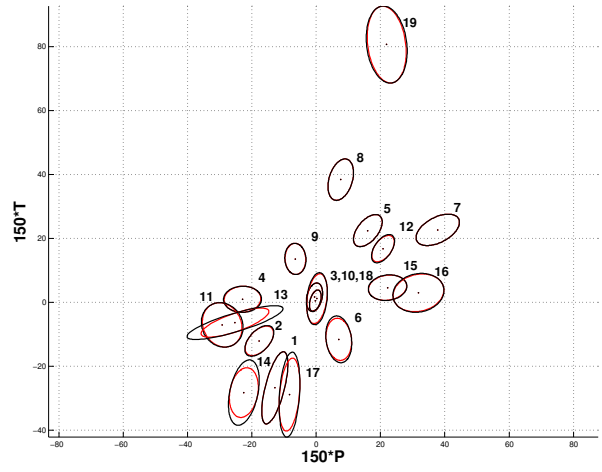


Figure 2.8: Equal color-difference ellipsoids of the RIT-DuPont datasets in IPT

2.4. Color Spaces, Color-Difference Formulas and Color Appearance Models

2.4.1. IPT

In 1998, a simple, uniform color space (or color appearance model), IPT, was derived by Ebner and Fairchild [Ebner, 1998b] based on three constant perceived hue datasets: the Munsell data set [Wyszecki, 1982], Hung and Berns' [Hung, 1995], and Ebner and

Fairchild's [Ebner, 1998a] datasets. The IPT space was designed to accurately predict hue without affecting other color appearance attributes. The invertible mapping between XYZ tristimulus values for 1931 2º color-matching functions with an illuminant D65 to IPT space consists of a 3-by-3 matrix, which convert the tristimulus values to cone response signals, followed by a non-linearity to simulate the post-processing of human visual system and another 3-by-3 matrix to achieve an opponent color response space. The transformation of IPT from tristimulus values are shown in Equations (2.9) – (2.11). In Equation (2.9), the tristimulus values for Y are in the range of 0 – 1.

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.4002 & 0.7075 & -0.0807 \\ -0.2280 & 1.1500 & 0.0612 \\ 0.0 & 0.0 & 0.9184 \end{bmatrix} \begin{bmatrix} X_{D65} \\ Y_{D65} \\ Z_{D65} \end{bmatrix} \tag{2.9}$$

$$\begin{aligned} L' &= L^{0.43}; L >= 0 \\ L' &= -(-L)^{0.43}; L < 0 \\ M' &= M^{0.43}; M >= 0 \\ M' &= -(-M)^{0.43}; M < 0 \\ S' &= S^{0.43}; S >= 0 \\ S' &= -(-S)^{0.43}; S < 0 \end{aligned} \tag{2.10}$$

$$\begin{bmatrix} I \\ P \\ T \end{bmatrix} = \begin{bmatrix} 0.4000 & 0.4000 & 0.2000 \\ 4.4550 & -4.8510 & 0.3960 \\ 0.8056 & 0.3572 & -1.1628 \end{bmatrix} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \tag{2.11}$$

The range of the lightness axis *I* is 0 to 1 and about -1 to 1 for chromatic axes *P* and *T*. The space has a similar color scale to CIELAB if one multiplies *I* by 100, and *P* and *T* by 150. The IPT space was proved as more uniform in perceived hue than color space CIELAB and color appearance model CIECAM97s and CIECAM02 [Ebner,

1998b]. In this research, IPT space was selected as the potential candidate to derive the color model for its easy implementation and good color appearance prediction.

2.4.2. CIEDE2000

CIEDE2000 is the latest CIE recommended color-difference formula based on CIELAB, shown in Equations (2.12) – (2.17) [CIE Publ.142, 2001], [CIE Publ. 15, 2004]. It was derived with an emphasis on the modeling of five aspects: chroma dependency, hue dependency, a rotation term in the blue area, lightness transformation and neutral gray modeling. The chroma dependency modeling in CIEDE2000 is similar to the modeling in CIE94 [Berns, 1993]. A hue-angle dependent, hue-difference weighting function was derived with four hue-difference datasets [Berns, 2002]: Luo-Rigg dataset and RIT-DuPont-Witt dataset derived by Maier from the BFD-P, RIT-DuPont and Witt datasets [Maier, 1992], Luo-Hue dataset [Luo, 1999] and Qiao's dataset [Qiao, 1998]. The near neutral area performance was benefitted from the modification of *a'* in Equation (2.14). With the rotation coefficient R_T, the blue area performance was improved. In addition, the formula has a lightness modification parameter S_L to adjust the lightness difference depending on lightness values. The order of the importance of these five corrections of CIEDE2000 is as follows: chroma, hue, rotation term, lightness and gray [Melgosa, 2004].

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2 + R_T\left(\frac{\Delta C'}{k_C S_C}\right)\left(\frac{\Delta H'}{k_H S_H}\right)} \qquad (2.12)$$

$$L' = L^* \qquad (2.13)$$

$$a' = (1+G)a^* \qquad (2.14)$$

$$b' = b^* \qquad (2.15)$$

$$C' = \sqrt{a'^2 + b'^2} \tag{2.16}$$

$$h' = \arctan\left(b'/a'\right) \tag{2.17}$$

where

$$G = 0.5\left(1 - \sqrt{\frac{\overline{C}_{ab}^{*\,7}}{\overline{C}_{ab}^{*\,7} + 25^7}}\right)$$

$$S_L = 1 + \frac{0.015\left(\overline{L}' - 50\right)^2}{\sqrt{20 + \left(\overline{L}' - 50\right)^2}}$$

$$S_C = 1 + 0.045\overline{C}'$$

$$S_H = 1 + 0.015\overline{C}'T$$

$$T = 1 - 0.17\cos(\overline{h}' - 30^\circ) + 0.24\cos(2\overline{h}') + 0.32\cos(3\overline{h}' + 6^\circ) - 0.20\cos(4\overline{h}' - 63^\circ)$$

$$R_T = -\sin(2\Delta\theta)R_C$$

$$\Delta\theta = 30\exp\left\{-\left[\left(\overline{h}' - 275^\circ\right)/25\right]^2\right\}$$

$$R_C = 2\sqrt{\frac{\overline{C}'^7}{\overline{C}'^7 + 25^7}}$$

2.4.3. CAM02 SCD LCD and UCS

The CAM02-SCD, CAM02-LCD and CAM02-UCS were introduced in 2006 [Luo, 2006]. They are color-difference formulas associating with uniform color spaces based on the best available color appearance model, CIECAM02. The compression on chroma and

transformation on lightness were applied on the colorimetric attributes *JMh* instead of *JCh* for better small color-difference modeling ability [Li, 2003]. The formula form shown in Equation (2.18) was optimized for small color-difference datasets, large color-difference datasets and their combination, separately. The optimized coefficients for each version of color-difference formulas are shown in Table 2.2.

$$\Delta E' = \sqrt{\left(\Delta J' / K_L\right)^2 + \Delta a'^2 + \Delta b'^2} \tag{2.18}$$

$$J' = \frac{(1+100c_1)J}{1+c_1 J} \tag{2.19}$$

$$M' = (1/c_2)\ln(1+c_2 M) \tag{2.20}$$

$$a' = M'\cos(h) \tag{2.21}$$

$$b' = M'\sin(h) \tag{2.22}$$

Table 2.2. Coefficients of CAM02-LCD, SCD and UCS [Luo, 2006].

Versions	CAM02-LCD	CAM02-SCD	CAM02-UCS
K_L	0.77	1.24	1.00
c_1	0.007	0.007	0.007
c_2	0.0053	0.0363	0.0228

2.4.4. CIECAM02 series

A series of Euclidized, CIECAM02-based color spaces were introduced by Berns and Xue in 2007 [Berns, 2007], [Xue 2008]. These color spaces were derived from the optimized color-difference formulas based on CIECAM02 *JCh* with a first-order polynomial compression on chroma and a second-order polynomial compression on lightness. The rotation matrix shown in Equation (2.24) can either be added in the color-difference formula and Euclidean color space or not, which resulted in two sets of color

models, each including one color-difference formula and corresponding Euclidean color space. Since the coefficients of these series color spaces were optimized only using the RIT-DuPont datasets, the color-difference formulas fit the RIT-DuPont dataset well, but not for other small color-difference datasets. This provided the clear concept of deriving Euclidean color spaces based on an optimized color-difference formula. The color-difference formulas are shown in Equations (2.23) – (2.24). The Euclidean color spaces are shown in Equations (2.25) – (2.30). The rotation matrix shown in Equation (2.24) can either be added on the color-difference formula and Euclidean color space or not, which resulted two sets of color models, each includes one color-difference formula and corresponding Euclidean color space. Since the coefficients of these series color spaces were optimized only using the RIT-DuPont datasets, the color-difference formulas fit the RIT-DuPont dataset well, but not for other small color-difference datasets. This provided the clear concept of deriving Euclidean color space based on an optimized color-difference formula.

$$\Delta E_{CAM,OPT} = \sqrt{\left(\frac{\Delta J'}{k_J S_J}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2} \tag{2.23}$$

where

$$\begin{pmatrix} J' \\ a' \\ b' \end{pmatrix} = \begin{pmatrix} 1.000 & 0.015 & -0.032 \\ -0.166 & 0.934 & -0.086 \\ -0.060 & -0.088 & 0.892 \end{pmatrix} \cdot \begin{pmatrix} J \\ a \\ b \end{pmatrix} \tag{2.24}$$

with $S_J = 0.5 + 1.25\left(\overline{J'}/100\right)^2$

$S_C = 1 + 0.02\overline{C'}$

$$S_H = 1 + 0.01\overline{C'}$$

$$\Delta E_{Euclidean} = \sqrt{\left(\Delta J^E\right)^2 + \left(\Delta a^E\right)^2 + \left(\Delta b^E\right)^2} \tag{2.25}$$

where

$$J^E = 0.99 \times \frac{200}{\sqrt{2.5}} \arctan \frac{\sqrt{2.5}}{100} J \tag{2.26}$$

$$C^E = 0.94 \times 50 \ln\left(1 + 0.02C\right) \tag{2.27}$$

$$h^E = h \tag{2.28}$$

$$a^E = C^E \cos\left(h^E\right) \tag{2.29}$$

$$b^E = C^E \sin\left(h^E\right) \tag{2.30}$$

2.4.5. DIN99, DIN99d, DIN99o

The DIN99 color-difference formula was developed in 1999 and later adopted as the German standard [DIN6176, 2000]. This formula applied a logarithmic transformation on the lightness channel and a rescaling of the CIELAB attributes L^* and C^*_{ab}, shown in Equations (2.31) – (2.39). Also, a rotation and stretch was applied on chroma plane and new chromatic attributes e and f were achieved in Equation (2.33) and (2.34). The Euclidean distance ΔE_{99} was adopted as the measure of color difference.

$$\Delta E_{99} = \frac{1}{k_E} \sqrt{\Delta L_{99}^2 + \Delta a_{99}^2 + \Delta b_{99}^2} \tag{2.31}$$

where

$$L_{99} = 105.51 \cdot \ln\left(1 + 0.0158L^*\right) \tag{2.32}$$

$$e = a^* \cos(16^\circ) + b^* \sin(16^\circ) \tag{2.33}$$

$$f = 0.7\left[b^* \cos(16^\circ) - a^* \sin(16^\circ)\right] \tag{2.34}$$

$$G = \sqrt{e^2 + f^2} \tag{2.35}$$

$$C_{99} = \ln(1 + 0.045G)/0.045 \tag{2.36}$$

$$h_{99} = \arctan\left(\frac{f}{e}\right) \tag{2.37}$$

$$a_{99} = C_{99} \cos(h_{99}) \tag{2.38}$$

$$b_{99} = C_{99} \sin(h_{99}) \tag{2.39}$$

A series of modifications were made on the DIN99 color space in 2002 and named as DIN99b, DIN99c and DIN99d [Cui, 2002]. Among them, DIN99d has the best performance and is shown in Equations (2.40) – (2.49). In the DIN99d formula, tristimulus values X was modified by subtracting a portion of Z to improve the performance in the blue area [Kuehni, 1999]. A different degree of rotation to DIN99 was applied in the new color space DIN99d.

$$\Delta E_{99d} = \frac{1}{k_E}\sqrt{\Delta L_{99d}^2 + \Delta a_{99d}^2 + \Delta b_{99d}^2} \tag{2.40}$$

$$X' = 1.12X - 0.12Z \tag{2.41}$$

$$L_{99d} = 325.22 \cdot \ln(1 + 0.0036L^*) \tag{2.42}$$

$$e = a^* \cos(50^\circ) + b^* \sin(50^\circ) \tag{2.43}$$

$$f = 1.14\left[-a^* \sin(50^\circ) + b^* \cos(50^\circ)\right] \tag{2.44}$$

$$G = \sqrt{e^2 + f^2} \tag{2.45}$$

$$C_{99d} = 22.5 \cdot \ln(1 + 0.06G) \tag{2.46}$$

$$h_{99d} = \arctan(f/e) + 50^\circ \tag{2.47}$$

$$a_{99d} = C_{99d} \cos(h_{99d}) \tag{2.48}$$

$$b_{99d} = C_{99d} \sin(h_{99d}) \tag{2.49}$$

DIN99o is the latest modification on DIN99 color space [Witt, 2009]. DIN99o has a similar form as DIN99b, but a parametric factor was added to the lightness channel for future use in compensating for differences between different datasets.

$$\Delta E_{99o} = \sqrt{\Delta L_{99o}^2 + \Delta a_{99o}^2 + \Delta b_{99o}^2} \tag{2.50}$$

$$L_{99o} = (1/k_E) \cdot 303.67 \cdot \ln(1 + 0.0039L^*) \tag{2.51}$$

$$e_o = a^* \cos(26^\circ) + b^* \sin(26^\circ) \tag{2.52}$$

$$f_o = -0.83 \cdot \left[-a^* \sin(26^\circ) + b^* \cos(26^\circ)\right] \tag{2.53}$$

$$G_o = \sqrt{e_o^{\,2} + f_o^{\,2}} \tag{2.54}$$

$$h_{eofo} = \arctan(f_o / e_o) \tag{2.55}$$

$$h_{99o} = h_{eofo} \cdot 180/\pi + 26^{\circ} \tag{2.56}$$

$$C_{99o} = \left(\frac{1}{0.0435 \cdot k_{CH} k_E} \right) \cdot \ln(1 + 0.075 G_o) \tag{2.57}$$

$$a_{99o} = C_{99o} \cos(h_{99o}) \tag{2.58}$$

$$b_{99o} = C_{99o} \sin(h_{99o}) \tag{2.59}$$

2.4.6. OSA UCS GP and OSA UCS E

OSA-UCS was an empirical color system for large color differences developed by the Optical Society of America's committee on Uniform Color Scales, led by MacAdam and published in 1974 [MacAdam, 1974]. Even though the space is non-uniform, the straight lines radiating from any color samples are geodesic lines with uniform scales [Huertas, 2006]. Thus, OSA-UCS was adopted as the basis of small color-difference formulas in 2006 [Huerta, 2006]. A formula similar to CIE94 was optimized for small color-difference datasets and named as ΔE_{GP}, shown in Equation (2.60).

$$\Delta E_{GP} = 10 \cdot \sqrt{\left(\frac{\Delta L_{OSA}}{k_L S_L} \right)^2 + \left(\frac{\Delta C_{OSA}}{k_C S_C} \right)^2 + \left(\frac{\Delta H_{OSA}}{k_H S_H} \right)^2} \tag{2.60}$$

where

$$S_L = 2.499 + 0.007\left(10\overline{L}_{OSA}\right)$$

$$S_C = 1.235 + 0.058\left(10\overline{C}_{OSA}\right)$$

$$S_H = 1.392 + 0.017\left(10\overline{C}_{OSA}\right)$$

In 2008, the OSA-UCS was modified and a new Euclidean color space was developed [Oleari, 2008]. An improvement was made to the formula based on integration along lightness and chroma directions. The Euclidean distance ΔE_E is applied as color-difference formula and shown in Equations (2.61) – (2.68).

$$\Delta E_E = \sqrt{(\Delta L_E)^2 + (\Delta G_E)^2 + (\Delta J_E)^2} \tag{2.61}$$

$$L_E = \left(\frac{1}{b_L}\right)\ln\left[1 + \frac{b_L}{a_L}(10L_{OSA})\right] \text{ with } a_L = 2.890,\ b_L = 0.015 \tag{2.62}$$

$$G_E = -C_E\cos(h) \tag{2.63}$$

$$J_E = C_E\sin(h) \tag{2.64}$$

$$h = \arctan\left(-\frac{J}{G}\right) \tag{2.65}$$

$$C_E = \left(\frac{1}{b_C}\right)\ln\left[1 + \frac{b_C}{a_C}(10C_{OSA})\right] \text{ with } a_C = 1.256,\ b_C = 0.050 \tag{2.67}$$

$$C_{OSA} = \sqrt{G^2 + J^2} \tag{2.68}$$

One difficulty to apply OSA-UCS based color spaces and color-difference formula is the complex transformation from tristimulus values to OSA-UCS *LGJ* colorimetric attributes. The recommended transformation was introduced by Oleari in 2004 [Oleari, 2004]. The transformation begins with tristimulus values under 1964 10° observer and CIE illumiant D65, and shown in Equations (2.69) – (2.73)

$$L_{OSA} = \frac{1}{\sqrt{2}}\left\{5.9\left[\left(Y_0^{1/3} - \frac{2}{3}\right) + 0.042 f(Y_0)\right] - 14.4\right\} \tag{2.69}$$

where

$$f(Y_0) = \begin{cases} (Y_0 - 30)^{1/3} & \text{if } Y_0 \geq 30 \\ -|Y_0 - 30|^{1/3} & \text{if } Y_0 < 30 \end{cases} \tag{2.70}$$

$$Y_0 = Y_{10}\left(4.4934 x_{10}^2 + 4.43034 y_{10}^2 - 4.2760 x_{10} y_{10} - 1.3744 x_{10} - 2.5643 y_{10} + 1.8103\right) \tag{2.71}$$

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{bmatrix} 0.6597 & 0.4492 & -0.1089 \\ -0.3053 & 1.2126 & 0.0927 \\ -0.0374 & 0.4795 & 0.5579 \end{bmatrix} \cdot \begin{pmatrix} X_{10} \\ Y_{10} \\ Z_{10} \end{pmatrix} \tag{2.72}$$

$$\begin{pmatrix} J \\ G \end{pmatrix} = \begin{bmatrix} S_J & 0 \\ 0 & S_G \end{bmatrix} \cdot \begin{bmatrix} -\sin\alpha & \cos\alpha \\ \sin\beta & -\cos\beta \end{bmatrix} \cdot \begin{pmatrix} \ln\left(\frac{A/B}{A_n/B_n}\right) \\ \ln\left(\frac{B/C}{B_n/C_n}\right) \end{pmatrix}$$

$$= \begin{bmatrix} 2(0.5735 L_{OSA} + 7.0892) & 0 \\ 0 & -2(0.7640 L_{OSA} + 9.2521) \end{bmatrix} \cdot \begin{bmatrix} 0.1792 & 0.9837 \\ 0.9482 & -0.3175 \end{bmatrix} \cdot \begin{pmatrix} \ln\left(\frac{A/B}{0.9366}\right) \\ \ln\left(\frac{B/C}{0.9807_n}\right) \end{pmatrix}$$

(2.73)

2.5. Color-Difference Spaces Based On Multi-Stage Color Vision Theory and Line Integration

In 2008, a series of generalized color-difference spaces having common structures as shown in Equation (2.74) – (2.79) were introduced by Berns based on the understanding

of human color vision [Berns, 2008]. These color-difference spaces have a similar transformation from tristimulus values to IPT space to model multi-stage color vision theory: a chromatic adaptation transformation ensures the color appearance property in the first step (Equation (2.74)); a linear transformation from tristimulus values to pseudo cone fundamentals is the second step to simulate the linear processing at the cones of human visual receptors (Equation (2.75)); a nonlinear stage was realized by an exponential function after the linear processing (Equation (2.76)); also, the opponent signals were generated by another linear transformation (Equation (2.77)). Besides the multi-stage color vision theory, finally, the line integration along chroma was adopted to compensate the increase of color tolerance with the increase of chroma (Equation (2.78)).

The Euclidean distance of the spaces was adopted as the measure of color difference, shown in Equation (2.79). The parameters in these models were optimized to achieve minimum deviation between visual color differences and calculated numerical color differences for the RIT-DuPont and Qiao datasets. The performance of these spaces as color-difference metric is equal or superior to CIEDE2000. These color-difference spaces have better interpolation on the understanding of human color vision than other color-difference formulas based on CIELAB.

$$\begin{pmatrix} X_{\text{Illuminant E}} \\ Y_{\text{Illuminant E}} \\ Z_{\text{Illuminant E}} \end{pmatrix} = M_{CAT02}^{-1} M_{VK} M_{CAT02} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{2.74}$$

$$\begin{pmatrix} L \\ M \\ S \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{pmatrix}_{cones} \begin{pmatrix} X_{\text{Illuminant E}} \\ Y_{\text{Illuminant E}} \\ Z_{\text{Illuminant E}} \end{pmatrix} \tag{2.75}$$

where $(e_1+e_2+e_3)=1$, $(e_4+e_5+e_6)=1$ and $(e_7+e_8+e_9)=1$

$$\begin{pmatrix} L' \\ M' \\ S' \end{pmatrix} = \begin{pmatrix} L^{1/\gamma} \\ M^{1/\gamma} \\ S^{1/\gamma} \end{pmatrix} \tag{2.76}$$

$$\begin{pmatrix} W' \Leftrightarrow K' \\ R' \Leftrightarrow G' \\ Y' \Leftrightarrow B' \end{pmatrix} = \begin{pmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 100 \end{pmatrix} \times \begin{pmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{pmatrix}_{opponency} \begin{pmatrix} L' \\ M' \\ S' \end{pmatrix} \tag{2.77}$$

where $(e_1+e_2+e_3)=1$, $(e_4+e_5+e_6)=1$ and $(e_7+e_8+e_9)=1$

$$\begin{pmatrix} L^E \\ a^E \\ b^E \end{pmatrix} = \begin{pmatrix} (W' \Leftrightarrow K') \\ (R' \Leftrightarrow G')f(C) \\ (Y' \Leftrightarrow B')f(C) \end{pmatrix} \tag{2.78}$$

where $f(C) = \dfrac{\ln(1+\beta_C \overline{C})}{\beta_C \overline{C}}$

and $C = \sqrt{(R' \Leftrightarrow G')^2 + (Y' \Leftrightarrow B')^2}$

$$\Delta E^E = \sqrt{(\Delta L^E)^2 + (\Delta a^E)^2 + (\Delta b^E)^2} \tag{2.79}$$

The coefficients of the color space 3 are illustrated in Table 2.3. The physiology was not considered in the color space 3 derivation and both the matrices were optimized only for the best STRESS value with pre-defined γ and β_C. The equal color-difference ellipsoids of the RIT-DuPont datasets were plotted in color space 3 in Figure. 2.9. Comparing to the ellipsoids of CIELAB shown in Figure 2.6, the rotation in blue region was greatly improved and the size of the ellipsoids are more uniform. In addition, the

constant hue datasets were plotted in Figure 2.10. In Figure 2.10, the color space 3 shows a good linearity similar to the IPT space. The pseudo cone fundamentals of color space 3 were calculated by the multiplication of the first matrix and color matching functions and shown in Figure 2.11. It is indicated by the great difference of the pseudo cone fundamentals to the color matching functions that the color matching and color difference measuring are in different levels of processing of human color vision.

Table 2.3. Color-space coefficients and STRESS values (part of Table I in [Berns, 2008])

Color Space	Matrix Cone	γ	Matrix Opponency	β_C	STRESS
3	$\begin{pmatrix} 1.2 & 0.2 & -0.4 \\ -0.7 & 1.1 & 0.6 \\ -0.6 & 0.9 & 0.7 \end{pmatrix}$	2.3	$\begin{pmatrix} 0.5 & 0.6 & -0.1 \\ 2 & -14 & 12 \\ 0 & 16 & -16 \end{pmatrix}$	0.04	20.6

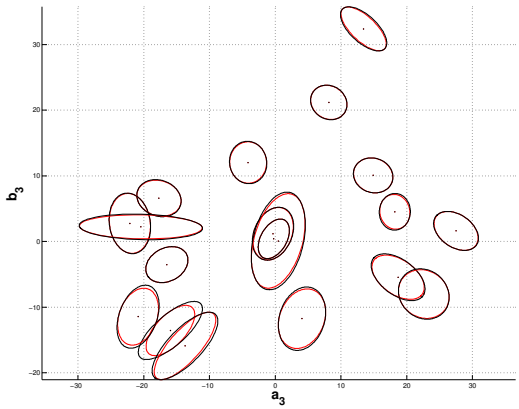


Figure 2.9. Equal color-difference ellipsoids in color space 3. [Berns, 2008]

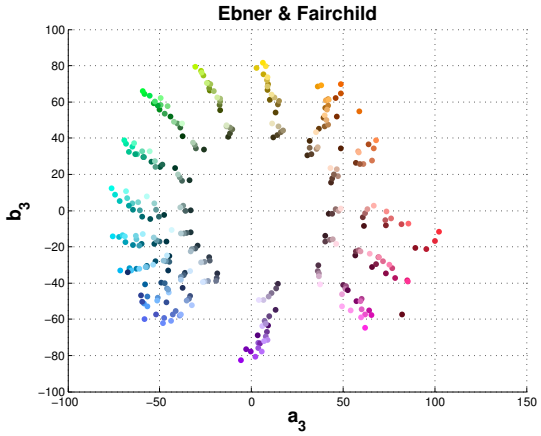


Figure 2.10. Ebner-Fairchild constant hue datasets in color space 3. [Berns, 2008]

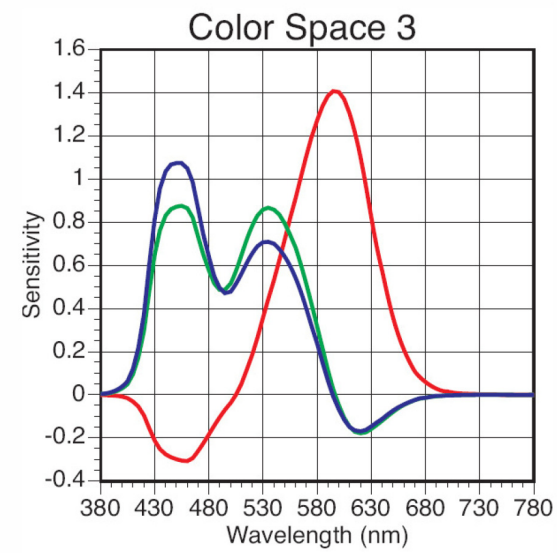


Figure 2.11. "Cone fundamentals" of color space 3. [Berns, 2008]

2.6. Derivation of Euclidean Color Spaces from Color-Difference Formulas with Analytical Method

The Euclidean color space can be developed by either analytical or computational methods to map the non-linear, non-uniform color spaces to linear and uniform color spaces based on the clues provided by different color-difference formulas optimized for color-difference datasets [Völz, 1998], [Völz, 1999-2000], [Thomsen, 2000], [Völz, 2006], [Urban, 2007a], [Urban, 2007b]. The analytical transformation of Euclidean color

space is desired since it provides clear physical meaning and is easy to implement. As an example, a Euclidean color space, L^{*}' a^{*}' b^{*}' derived based on CIE94 is shown in Equations (2.80) – (2.86) [Thomsen, 2000]

$$L^{*}' = L^{*} \tag{2.80}$$

$$a^{*}' = a^{*} f\left(C^{*}\right) \tag{2.81}$$

$$b^{*}' = b^{*} f\left(C^{*}\right) \tag{2.82}$$

$$C^{*} = \sqrt{a^{*2} + b^{*2}} \tag{2.83}$$

$$f\left(C^{*}\right) = \frac{C^{*}'}{C^{*}} = \frac{\ln\left(1 + 0.045C^{*}\right)}{0.045C^{*}} \tag{2.84}$$

The compression along chroma is derived from the following integration:

$$dC^{*}' = dC^{*} / \left(1 + 0.045C^{*}\right) \tag{2.85}$$

$$C^{*}' = \int_{0}^{C^{*}} dt / \left(1 + 0.045t\right) = \frac{\ln\left(1 + 0.045C^{*}\right)}{0.045} \tag{2.86}$$

Here, the compression is only applied in the chroma direction since it is difficult to achieve exact integration along hue direction. Usually, the form of Equation (2.86) is kept and the coefficients P and Q in Equation (2.87) were optimized again to achieve better agreement between the Euclidean distance in the new space and the color difference calculated with CIE94.

$$C^{*}' = \int_{0}^{C^{*}} dt / \left(1 + 0.045t\right) = \frac{\ln\left(1 + PC^{*}\right)}{Q} \tag{2.87}$$

In addition, randomized positions in color space were sampled. For each position, several random direction and random distance within the range of small color difference were generated as batches to evaluate the agreement [Thomsen, 2000], [Berns, 2007]. This analytical method of deriving Euclidean color space from optimized color-difference formulas was adopted in this research.

3. Evaluating Color Difference Equation Performance Incorporating Visual Uncertainty

Note that this Chapter was submitted as a revised final article to Color Research and Application, authored by S. Shen and R. S. Berns, respectively. It has been accepted and is in press.

3.1. Abstract

5000 randomized ellipsoids for each of 19 color centers comprising the RIT-DuPont dataset were generated based on both the tolerance median (T50) and visual uncertainty (fiducial limits). When plotted as two-space projections, they provided a qualitative description of the ellipsoid reliability, this reliability dependent on visual uncertainty. The ellipsoids were considered as local color-difference equations. STRESS was calculated for each ellipsoid, quantifying the deviation between visual color differences and numerical color differences calculated by the T50-based ellipsoids, the randomized ellipsoids and three color-difference equations: CIELAB, CIE94, and CIEDE2000. F-tests comparing STRESS determined statistical significance between calculated and visual color differences. The percentage of randomized ellipsoids that were beyond the critical *F* values was used as a metric for determining whether a color difference equation was under- or over-fitting the visual data. A non-ellipsoid method and an average standard error method were developed and tested for cases when the dataset may not enable ellipsoid fitting and where uncertainty has been reported only as an average standard error. In the latter case, only equation under-fitting could be determined. Thus visual uncertainty can be used as a criterion in both equation development and evaluation.

3.2. Introduction

During the late 1980's two visual experiments were performed to develop a suprathreshold color-difference dataset to be used for improving the relationship between visual and numerical color difference evaluation [Reniff, 1989], [Snyder, 1991], [Balonon-Rosen, 1993]. 957 color-difference pairs produced using a glossy automotive paint system were compared to a single near-neutral anchor pair with color difference of 1.02 ΔE^*_{ab} and judged by two sets of 50 observers. Based on the results of these experiments, a combined dataset with 19 color centers was developed, which was designated as the RIT-DuPont dataset [Berns, 1991]. The dataset was composed of 156 color-difference vectors about 19 color centers. For each color center, 6 – 14 color-difference vectors directionally sampled CIELAB space. The length between the origin and the median tolerance (T50) of each color-difference vector represented the equal visual color difference to the anchor pair. The 95% confidence limit of each color-difference vector was described by an LFL and a UFL (lower and upper fiducial limits), which provided uncertainty information such as how well the observational data fit a cumulative normal distribution and the observational standard deviation. These fiducial limits correspond to a standard error of ±5.5%, considerably smaller than other visual color discrimination experiments, e.g. ±8.9% for the Luo-Rigg dataset [Luo, 1986], ±10.0% for the Cheung-Rigg dataset [Cheung, 1986], ±11.5% for the Qiao, et al. dataset [Qiao, 1998], and ±8.7% for the Kim, et al. dataset [Kim, 2001]. This calculation is described below in detail. It differs from the ±3.5%, reported previously, based on the average standard deviation and the number of observers [Berns, 1991]

The RIT-DuPont dataset was used in the derivation of CIE94 and CIEDE2000 color-difference equations [Berns, 1993], [Luo, 2001]. In deriving the CIE94 color-difference equation, a weighting factor was added onto each color-difference vector, which was based on their visual uncertainty [Berns, 1993]. In 1997, Melgosa, et al. performed an analysis based on the RIT-DuPont dataset to fit contours of equal color difference (ellipsoids) at the 19 color centers in CIELAB and $x,y,Y/100$ spaces [Melgosa, 1997]. The T50 positions of the color-difference vectors of each color center were used in fitting these ellipsoids. However, because the number of directional sampling was quite small (6 – 14), it seems that the shape and the orientation of the ellipsoids might be sensitive to slight changes of the color-difference vectors due to observer uncertainty. This could have a dramatic effect when using the RIT-DuPont dataset for equation development and testing. Thus, it was of interest to evaluate the stability of these equal color-difference contours considering visual uncertainty and develop a methodology to incorporate visual uncertainty into equation testing.

3.3. Ellipsoids Fitting Procedures

The basic discrimination ellipsoid fitting procedures were well described by Indow and Morrison [Indow, 1991]. Melgosa, et al. provided the detailed procedures of fitting the RIT-DuPont dataset in CIELAB [Melgosa, 1997]. The six metric coefficients characterizing each ellipsoid were obtained by minimizing the objective function Z, defined as:

$$Z = \sum_i \left(\Delta V_i^2 - e_i^2 \right) \tag{3.1}$$

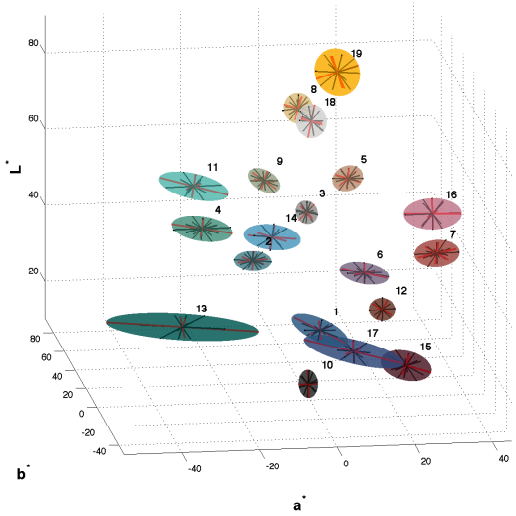
where ΔV is the visual color difference, constant for all the color pairs in the RIT-DuPont dataset, i represents the different color pairs, and e^2 represents the ellipsoid defined in CIELAB space as:

$$e_i^2 = b_{11}\left(\Delta a_i^*\right)^2 + b_{22}\left(\Delta b_i^*\right)^2 + 2b_{12}\left(\Delta a_i^*\right)\left(\Delta b_i^*\right) + 2b_{13}\left(\Delta a_i^*\right)\left(\Delta L_i^*\right) + 2b_{23}\left(\Delta b_i^*\right)\left(\Delta L_i^*\right) + b_{33}\left(\Delta L_i^*\right)^2 \tag{3.2}$$

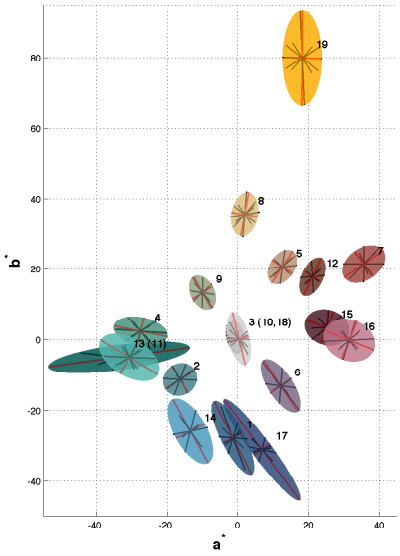
For one color center, each of 6 – 14 color-difference vectors sampling CIELAB space was used in calculating one e^2, that is, estimating the model coefficients b_{11}, b_{22},…,b_{33}. The T50 point of each color-difference vector represented the length of this color-difference vector. A constant visual color difference $\Delta V = 1.0$ from the origin to each of the T50 points was used because in the RIT-DuPont experiment, the anchor pair was the same for all the observations. Because of the method used to prepare samples along each vector direction, the actual origins of the color-difference vectors were slightly different to each other. The average origin of all color-difference vectors of one color center was used to represent the origins of all these vectors, and all the color-difference vectors were shifted according to this average origin. Totally, there were 156 color-difference vectors for 19 color centers (listed in the Table IV in the article of Berns et al [Berns, 1991]). All but four of them were used in fitting 19 ellipsoids. The fiducial limits of these four omitted color-difference vectors had values of negative or positive infinity [Berns, 1991], [Melgosa, 1997]. For these 152 color-difference vectors, the maximum origin shift was 2.72 ΔE^*_{ab} and the average origin shift was 0.67 ΔE^*_{ab}. Since the RIT-DuPont experiment assumed symmetry of the color-difference vectors in CIELAB space, both the provided median tolerances (+T50) of the color-difference vectors and their symmetric

points (-T50) with respect to the origins were used in fitting the ellipsoids. These optimizations were repeated in this research to validate the methodology. In this case *fminunc*, an unconstrained nonlinear optimization function in the Matlab optimization toolbox, was used. The model coefficients of these ellipsoids agreed well with the results in the article of Melgosa et al. [Melgosa, 1997] with differences smaller than 0.0001, verifying the methodology.

The T50-based ellipsoids for the 19 color centers comprising the RIT-DuPont dataset are plotted in Figure 3.1. The ellipsoids were enlarged three times their actual size to show them clearly. The black lines represent the visual color-difference vectors of the RIT-DuPont data. The red lines represent the principal axes of the ellipsoids. In Figure 3.1 (b) (c) and (d), the ellipsoids are plotted as projections onto each listed two-dimensional plane. Ordinarily, these are plotted as cross-sections. Projections were used to better show the relationship between the visual color-difference vectors and the fitted ellipsoids. A Matlab toolbox, *Ellipsoidal Toolbox*, was used in drawing the ellipsoids [Kurzhanskiy, 2006].

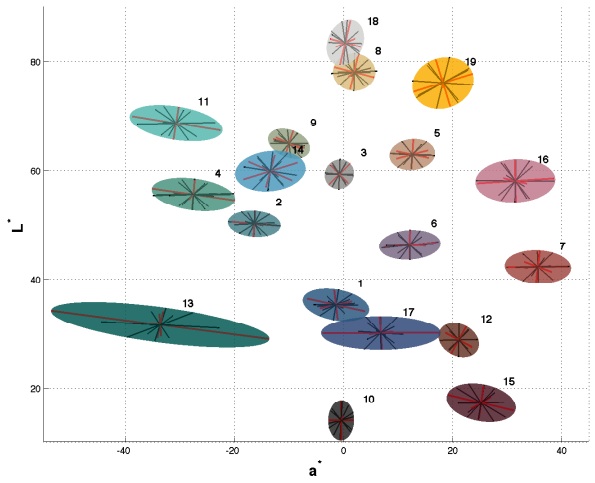


(a) Three-dimensional view.

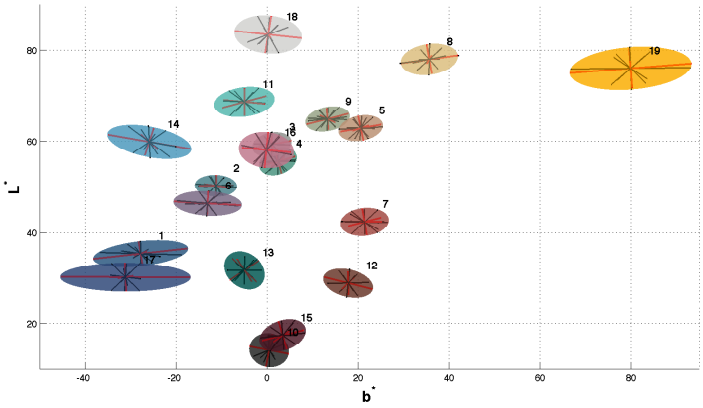


(b) Projection onto the *a*b** plane.

Figure 3.1. Equal color difference contours of the RIT-DuPont dataset (part 1 of 2).



(c) Projection onto the a^*L^* plane.



(d) Projection onto the b^*L^* plane.

Figure 3.1. Equal color difference contours of the RIT-DuPont dataset (continued, part 2 of 2).

The radius of a sphere having the same volume as the fitted ellipsoid was calculated for each color center. The radius has units of ΔE^*_{ab} and provides an intuitive metric approximating the ellipsoid. The relationship between the average length of the color-difference vectors and the radii are shown in Figure 3.2.

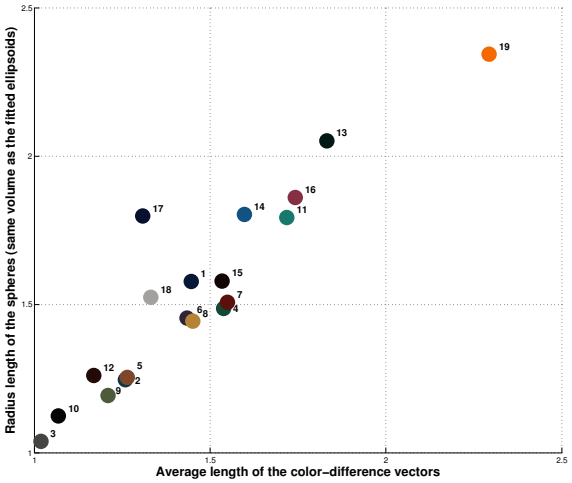
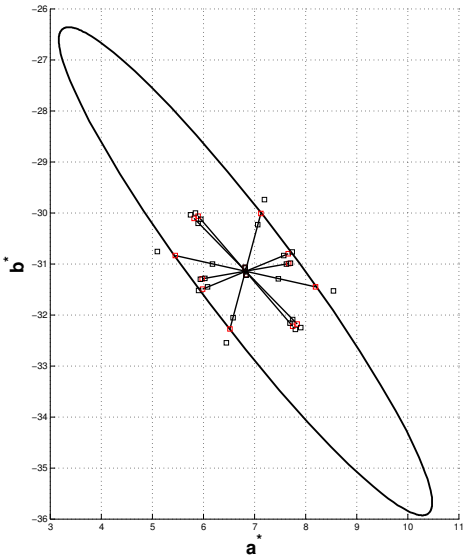


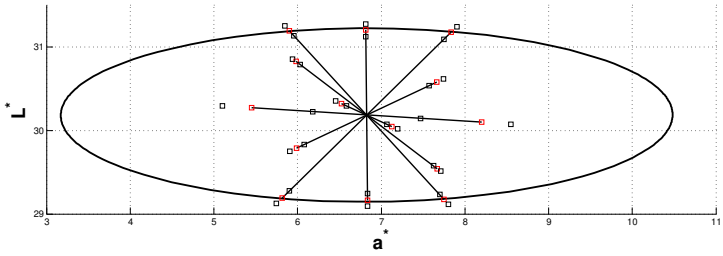
Figure 3.2. Correlation between the T50-based ellipsoids and the visual color-difference vectors of the RIT-DuPont dataset.

An approximately linear relationship indicates good agreement between the fitted ellipsoids and the color-difference vectors. Color center 17 was an outlier, warranting a detailed analysis; the fitted ellipsoid, its principal axes and the color-difference vectors (T50, UFL and LFL) are shown in Figure 3.3. In Figure 3.3, the thin lines illustrate the visual color-difference vectors; red squares illustrate the T50 and black squares illustrate the UFL and the LFL. It was found that there was not a tolerance vector having a similar direction as the longest principal axis of the T50-based ellipsoid based on the observation on the 3-dimension plot from different directions. (As two-dimensional projections, this observation is not seen.) This particular set of T50 values and directions resulted in an ellipsoid that represented the experimental data well, but did not represent the actual equal color-difference contour due to the lack of tolerance vectors in the principal axis direction. The equal color-difference ellipsoids were good visualization tools to illustrate the color-different dataset. But if the volume of the ellipsoids is far beyond the volume suggested by the average lengths of color vectors, such as No. 17 in Figure 3.2, using

such ellipsoids to evaluate color space visual uniformity and equation performance will be misleading. Furthermore, it is possible that small differences in the T50s caused by visual uncertainty may change the orientation and volume of this fitted ellipsoid dramatically [Luo, 1986], [Alder, 1981], [Witt, 1987].

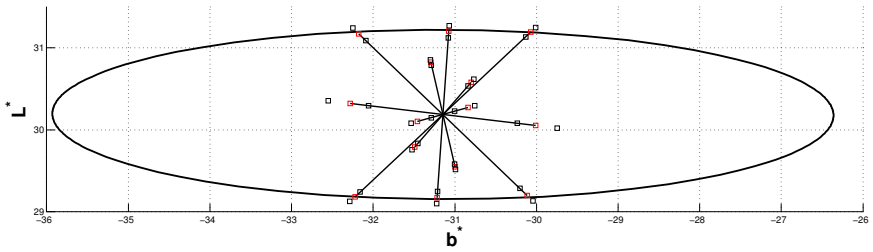


(a) Projection onto the a^*b^* plane.



(b) Projection onto the a^*L^* plane.

Figure 3.3. The T50-based ellipsoid of color center 17 of the RIT-DuPont dataset (part 1 of 2).



(c) Projection onto the b^*L^* plane.

Figure 3.3. The T50-based ellipsoid of color center 17 of the RIT-DuPont dataset (continued, part 2 of 2).

3.4. Generating Ellipsoids Considering Visual Uncertainty

The main technique used previously to evaluate the reliability of chromaticity-discrimination ellipses (in two-dimensional planes of Δx , Δy and ΔY space) was developed by Alder in 1981 [Alder, 1981]. In his method, the random deviates of color-difference data were produced using the Monte-Carlo method with an assumed normal distribution. Randomized ellipses were generated based on these new color-difference data. Then, the change of the randomized ellipses was evaluated. Luo and Rigg [Luo, 1986] used this technique to remove data with large uncertainty during the combining of different datasets and Cheung and Rigg [Cheung, 1986] and Witt [Witt, 1987] used the method to evaluate their datasets.

Here, a similar Monte-Carlo method was used to generate randomized ellipsoids. In this case the distribution of the random deviates of the color-difference data was determined by the uncertainty provided by the RIT-DuPont dataset rather than an estimated normal distribution. The randomized color-difference vectors corresponding to one original color-difference vector were along the direction of the vector, and their length followed the distribution determined by the UFL and the LFL of the vector. Since

the UFL and the LFL were the 95% confidence limit with the assumption that the frequency-of-rejection data in the visual experiment had a cumulative-normal distribution, the length distribution conformed to the following rules: The probabilities of the vector length appearing longer than T50 and shorter than T50 were the same (50% to 50%). The probabilities of the vector length appearing longer than T50 had a normal distribution with $\mu = T50$ and $\sigma = \left(\frac{UFL - T50}{2} \right)$; the division was required since a 95% confidence limit is equal to twice the standard error. The probabilities of the vector length appearing shorter than the T50 had a normal distribution with $\mu = T50$ and $\sigma = \left(\frac{T50 - LFL}{2} \right)$. Following these rules, 10,000 sets of color-difference vectors were generated randomly for each color center. The frequency of the lengths of the 10,000 randomized color-difference vectors for the seven color vectors of color center 17 is illustrated in Figure 3.4, in which the solid black lines illustrate T50 and the dashed black lines illustrate the UFL and the LFL. Because the UFL and the LFL were, by definition, asymmetric with respect to the T50, the length distributions were not a symmetric normal distribution.

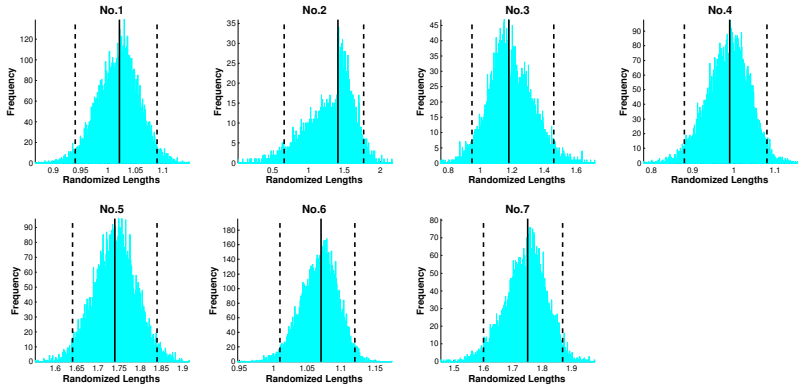
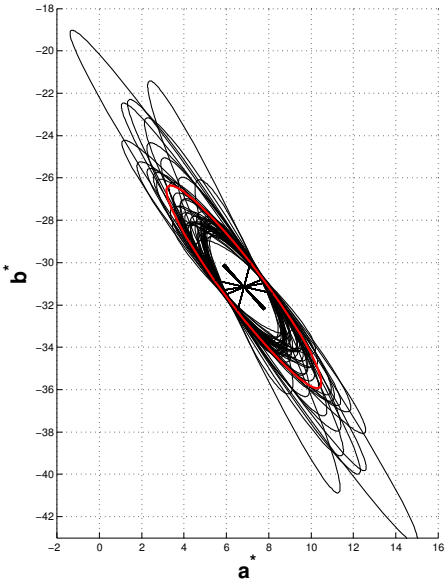


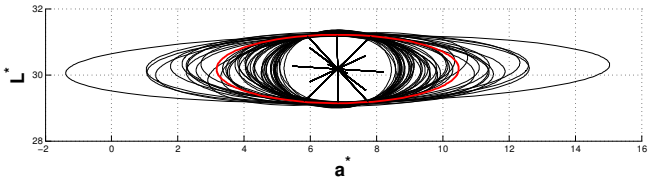
Figure 3.4. Vector length distribution of color center 17.

The lengths of the new generated color-difference vectors were randomly chosen from the length distribution. New ellipsoids were fit based on these randomized color-difference vectors. In some cases, a set of the randomized color-difference vectors could not be fit as an ellipsoid; that is, the optimized model coefficients did not correspond to an ellipsoid and they led to negative values for e^2. The same situation was also reported in previous research [Luo, 1986]. For cases where this occurred in the current research, the set was discarded. The process was repeated until 5,000 ellipsoids were generated for each color center.

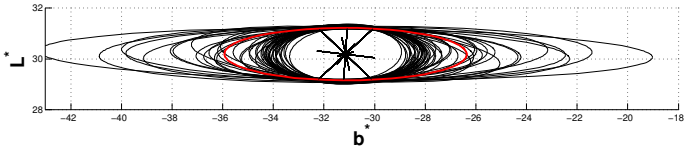
Fifty ellipsoids for color center 17 (dark blue), selected randomly from the 5,000 randomized ellipsoids, were compared with the T50-based ellipsoid and illustrated in Figure 3.5. In Figure 3.5, red ellipses are the projections of the T50-based ellipsoids, and the black lines illustrate projections of the visual color-difference vectors. In addition, axes are equally scaled in order to show the actual shape of the ellipses.



(a) Projections onto the *a*b** plane.



(b) Projections onto the *a*L** plane.

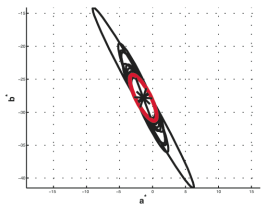


(c) Projections onto the *b*L** plane.

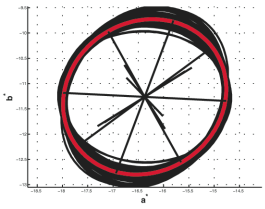
Figure 3.5. Fifty randomized ellipsoids for color center 17.

It is observed that the orientation was quite stable whereas the volume varied considerably. That is, the length of the major axis was very sensitive to visual uncertainty, supporting our previous analysis comparing the T50 vector lengths with its fitted ellipsoid. The randomized ellipsoids of color center 17 had large shape variance but small orientation variance. The projections of the 50 ellipsoids, randomly selected between the

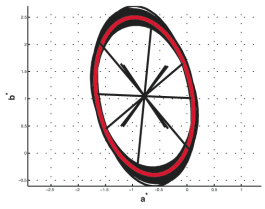
5000 ellipsoids generated, onto the a^*b^*, a^*L^* and b^*L^* planes for the other color centers are shown in Figure 3.6. Similar to Figure 3.5, red ellipses are the projections of the T50-based ellipsoids, and the black straight lines illustrate projections of the visual color-difference vectors. For a given color center, two axes (a^*b^*, a^*L^* or b^*L^*) are equally scaled in order to show the actual shape of the ellipses. Each subplot has been separately scaled to better visualize the variation in size and orientation. Large orientation variance was found for color centers 7, 9, 13, 14, 15, 16, 18 and 19. For color centers 1, 7, 11, 13, 14, 15, 16, 17, 18 and 19, large shape variance was found. The randomized ellipsoids for color centers 5, 6 and 8 had very small shape and orientation variance, because of the lower visual uncertainty of these color centers [Berns, 1991]. The randomized ellipsoid clouds at the three of five color centers recommended by the CIE [Robertson, 1978] (No. 1 blue, No. 3 gray, No. 4 green, No. 7 red and No.8 yellow) had quite small variance (except for blue and red).



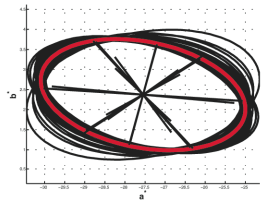
No. 1: Moderate blue



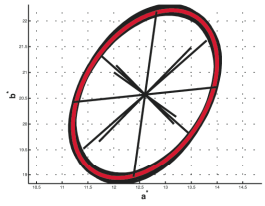
No. 2: Moderate greenish blue



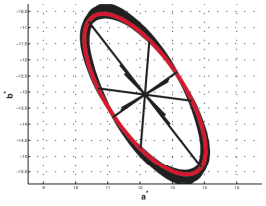
No. 3: Medium gray



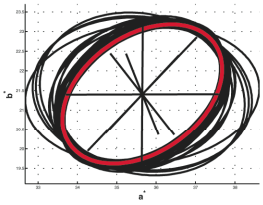
No. 4: Moderate bluish green



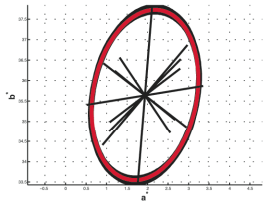
No. 5: Light brown



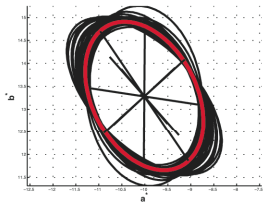
No. 6: Grayish purple



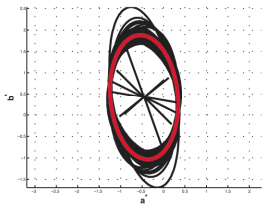
No. 7: Dark reddish orange



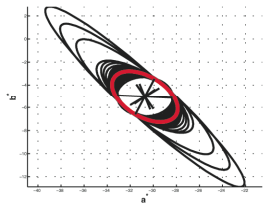
No. 8: Moderate yellow



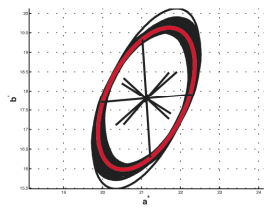
No. 9: Grayish yellow green



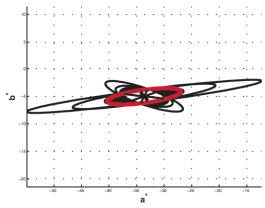
No. 10: Black



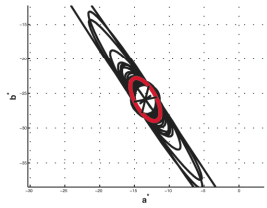
No. 11: Light bluish green



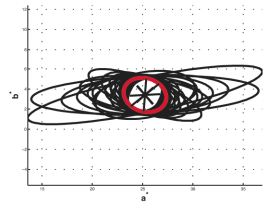
No. 12: Moderate reddish brown



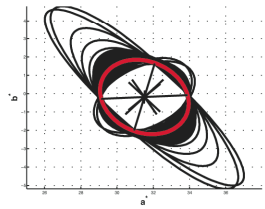
No. 13: Dark bluish green



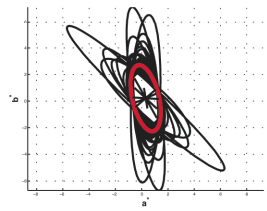
No. 14: Brilliant greenish blue



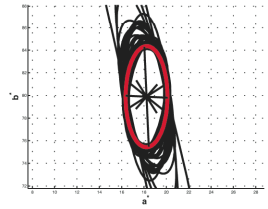
No. 15: Very dark red



No. 16: Moderate purplish pink

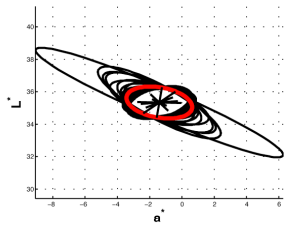


No. 18: Light gray

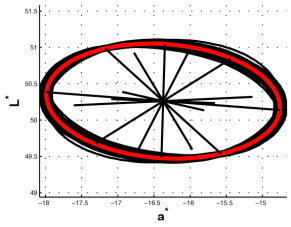


No. 19: Strong orange yellow

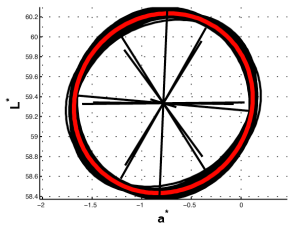
(a): *a*b** plane



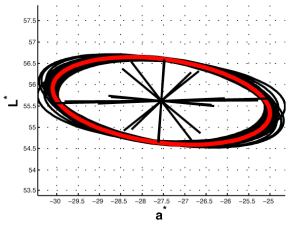
No. 1: Moderate blue



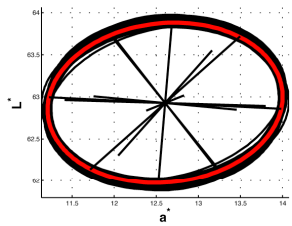
No. 2: Moderate greenish blue



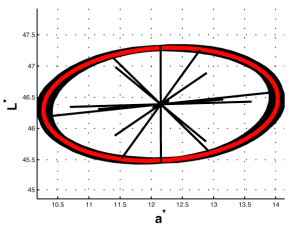
No. 3: Medium gray



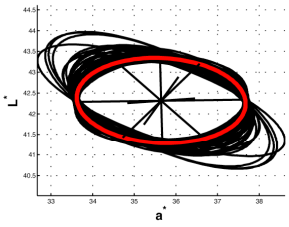
No. 4: Moderate bluish green



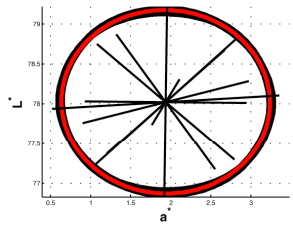
No. 5: Light brown



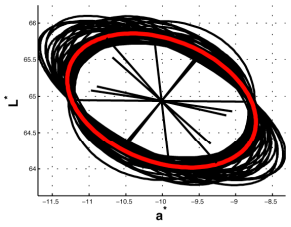
No. 6: Grayish purple



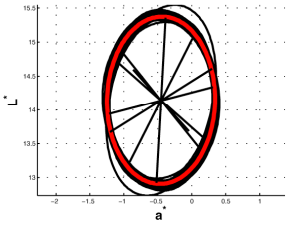
No. 7: Dark reddish orange



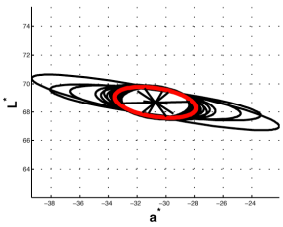
No. 8: Moderate yellow



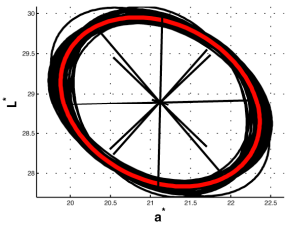
No. 9: Grayish yellow green



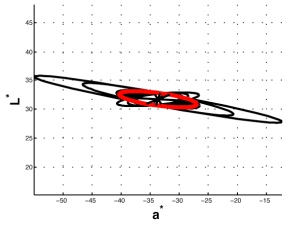
No. 10: Black



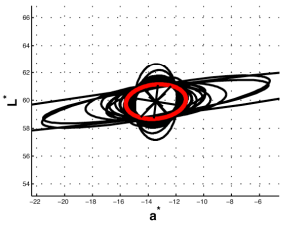
No. 11: Light bluish green



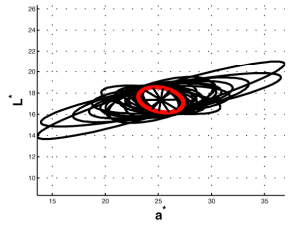
No. 12: Moderate reddish brown



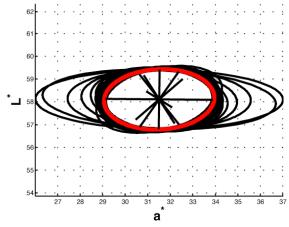
No. 13: Dark bluish green



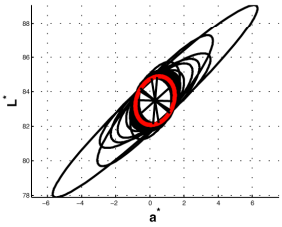
No. 14: Brilliant greenish blue



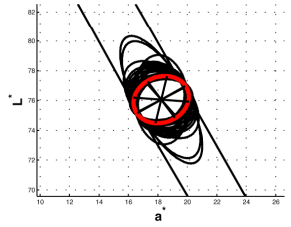
No. 15: Very dark red



No. 16: Moderate purplish pink

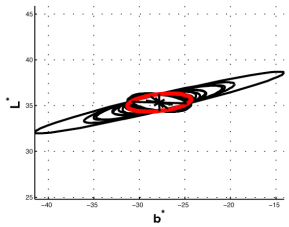


No. 18: Light gray

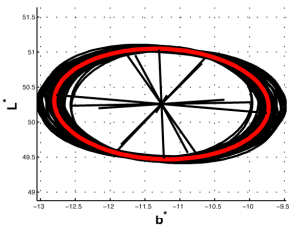


No. 19: Strong orange yellow

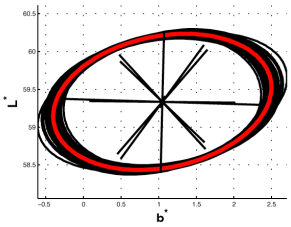
(b): a^*L^* plane



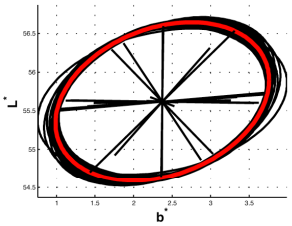
No. 1: Moderate blue



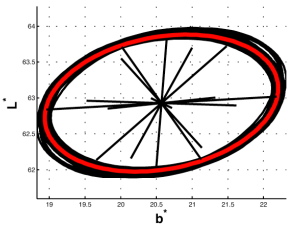
No. 2: Moderate greenish blue



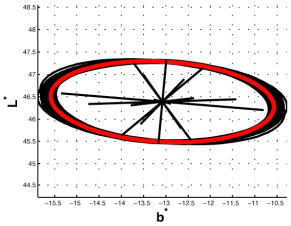
No. 3: Medium gray



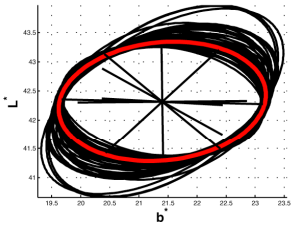
No. 4: Moderate bluish green



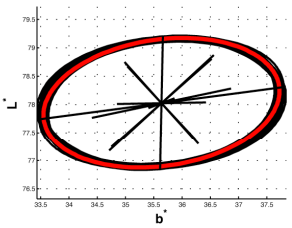
No. 5: Light brown



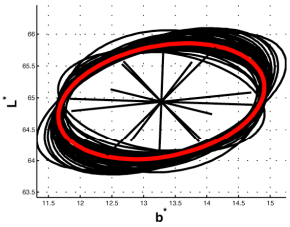
No. 6: Grayish purple



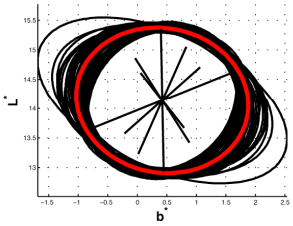
No. 7: Dark reddish orange



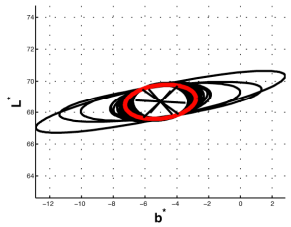
No. 8: Moderate yellow



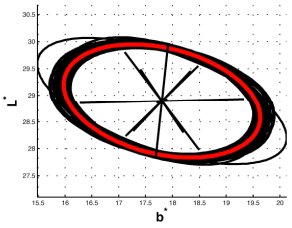
No. 9: Grayish yellow green



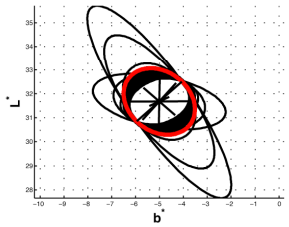
No. 10: Black



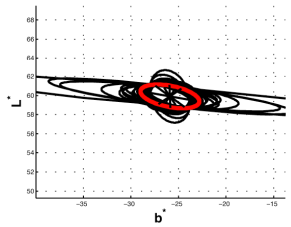
No. 11: Light bluish green



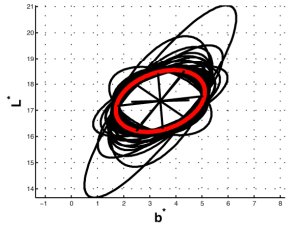
No. 12: Moderate reddish brown



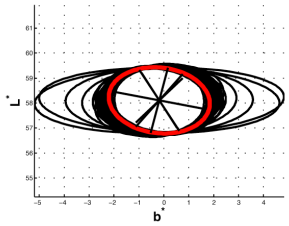
No. 13: Dark bluish green



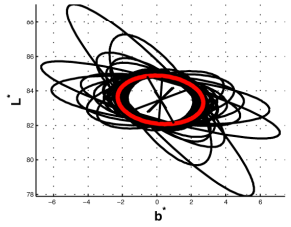
No. 14: Brilliant greenish blue



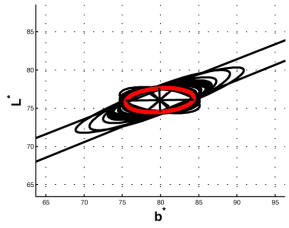
No. 15: Very dark red



No. 16: Moderate purplish pink



No. 18: Light gray



No. 19: Strong orange yellow

(c): b^*L^* plane

Figure 3.6. 50 randomized ellipsoids of the RIT-DuPont dataset.

3.5. Evaluation of Ellipsoid Variability

The principal axes of the ellipsoids provided sufficient shape (size) and orientation information to characterize an ellipsoid [Witt, 1999] and analyze variability in terms of the length and orientation.

3.5.1. Shape Variance

The lengths of the three principal axes of each of the 5,000 ellipsoids for each color center were calculated to characterize the shape and size of the randomized ellipsoids. The statistics of the lengths of the principal axes of the randomized ellipsoids are listed in Table 3.1. The three principal axes of each ellipsoid were designated as the 1st, 2nd, and 3rd principal axes by the order of their length: the longest principal axis was designated as the 1st principal axis, and the shortest was designated as the 3rd principal axis. A coefficient of variation (CV) was calculated to represent the normalized standard deviation for an axis. The average CV of all three principal axes is a qualitative metric defining the uncertainty for each color center. Color center 10 – 19 had larger uncertainty than color centers 1 – 9, which agreed with one of the conclusions given by Berns [Berns, 1991]: phase II of the RIT-DuPont experiment had larger observer variance than phase I of the experiment. Color center 14 and17 showed the largest variance of the principal axes.

Table 3.1. Statistics of the lengths of the principal axes of the randomized ellipsoids.

Color Center	Average of Principal Axes Length			Standard Deviation of Principal Axes Length			Average CV (%) of principal axes length for all three principle axes
	1st	2nd	3rd	1st	2nd	3rd	
1	4.36	1.11	0.90	2.70	0.02	0.04	22.66
2	1.71	1.47	0.78	0.08	0.06	0.02	4.00
3	1.50	0.93	0.81	0.07	0.03	0.02	3.20
4	2.63	1.36	0.93	0.08	0.07	0.02	3.30
5	1.82	1.19	0.92	0.04	0.04	0.04	3.18
6	2.94	1.18	0.90	0.08	0.02	0.02	2.19
7	2.45	1.63	0.98	0.93	0.22	0.06	19.16
8	2.13	1.28	1.11	0.04	0.03	0.03	2.05
9	1.81	1.18	0.82	0.12	0.04	0.06	5.78
10	1.54	1.21	0.79	0.19	0.04	0.02	6.15
11	3.72	1.78	1.03	2.54	0.04	0.07	25.65
12	2.03	1.04	0.96	0.17	0.03	0.03	4.60
13	5.03	1.33	0.99	3.40	0.06	0.04	25.53
14	3.72	1.65	1.19	4.58	0.14	0.19	49.25
15	3.24	1.78	1.03	2.97	0.19	0.05	35.88
16	2.87	1.92	1.31	1.39	0.10	0.04	18.73
17	6.24	1.08	0.91	8.03	0.05	0.04	45.89
18	3.52	1.43	0.96	3.86	0.13	0.06	41.62
19	6.39	1.93	1.49	7.45	0.11	0.09	42.80

3.5.2. Orientation Variance

The orientation of the ellipsoids was characterized by the orientation of the principal axes. Rich and Billmeyer developed a statistical test for a significant difference in orientation between two fitted chromaticity-discrimination color ellipses in 1983 [Rich, 1983]. Here, the analyses focused on the range of the variance of the randomized ellipsoid orientation instead of the difference between them. The orientation of the randomized ellipsoids was compared with the orientation of the T50-based ellipsoids for each color center. Since both the shape and the orientation of the randomized ellipsoids were different from each other, it was difficult to determine which principal axis of a randomized ellipsoid should be chosen and compared with the first principal axis of the T50-based ellipsoid. Thus, the angles between all three principal axes of the randomized ellipsoids and the first principal axis of the T50-based ellipsoid were calculated. The smallest in these three calculated angles was chosen to represent the directional deviation between the

randomized ellipsoid and the first principal axis of the T50-based ellipsoid. The same calculation was repeated for the second and the third principal axes. The average results of the calculation for all 5000 ellipsoids with unit of degree are shown in Table 3.2. There was quite large orientation variance of the randomized ellipsoids for color centers 2, 7, 9, 10, 13 – 16, 18 and 19. The plots in Figure 3.6 for 50 ellipsoids well illustrated the results for all 5,000 randomized ellipsoids. For the remaining color centers, the directional variance was small. Large orientation variance reflected the larger uncertainties of the visual data.

Table 3.2. Orientation variance of the randomized ellipsoids.

Color Center	Minimum of the three angles between three principal axes of the randomized ellipsoids and one of the principal axes of the T50-based ellipsoids (average for 5000 ellipsoids, degrees)		
	Comparing with the first Principal Axis of the T50-based ellipsoids	Comparing with the second Principal Axis of the T50-based ellipsoids	Comparing with the third Principal Axis of the T50-based ellipsoids
1	5.2	6.8	8.1
2	19.6	20.2	3.9
3	3.1	6.3	6.1
4	3.6	4.2	2.4
5	3.4	7.5	7.1
6	1.8	3.3	2.9
7	8.1	12.9	10.1
8	2.1	4.7	4.8
9	8.9	13.5	12.7
10	14.1	14.0	4.0
11	6.6	7.3	3.5
12	3.5	17.9	18.2
13	8.8	12.1	14.1
14	12.3	17.9	16.9
15	18.7	21.3	12.4
16	18.1	19.1	8.2
17	4.9	12.7	14.0
18	12.7	14.4	11.4
19	8.3	13.9	15.5

3.6. Using the Randomized Ellipsoids for Performance Evaluation

The variance of the randomized ellipsoids indicated a possible application for evaluating different color-difference equations: It is not required that an equation should predict visual color-differences or T50-based ellipsoids perfectly; rather, the prediction should have equivalent performance to the cloud of the randomized ellipsoids [Witt, 1987]. The

deviation between the randomized ellipsoids and the RIT-DuPont visual color-difference vectors represented the visual uncertainty range of the RIT-DuPont dataset. Therefore, this deviation provided a tool to evaluate color-difference equations and performance criteria when deriving new equations.

In 2007, a statistical index named standardized residual sum of squares (STRESS), which is a multidimensional scaling technique, was recommended to evaluate the performance of different color-difference equations [Garcia, 2007], [Melgosa, 2008]. The definition of STRESS for measuring the deviation between visual color differences and numerical color differences is:

$$STRESS = 100 \cdot \left(\frac{\sum_i \left(\Delta E_i - F \Delta V_i \right)^2}{\sum_i F^2 \Delta V_i^2} \right)^{1/2}$$

$$\text{with } F = \frac{\sum_i \Delta E_i^2}{\sum_i \Delta E_i \Delta V_i} \tag{3.4}$$

where ΔE defines the numerical color difference, ΔV defines the visual color difference, F is a scaling factor for adjusting ΔV values to the same scale as ΔE, and i is the number of color-difference pairs.

STRESS has the advantage of making statistical inferences compared with other metrics that have been developed in the past such as PF [Luo, 1987b], PF/3 [Guan, 1999], and others included in these two comprehensive evaluating factors: $V_{A\text{-}B}$ [Strocka, 1983], γ, CV, and a correlation coefficient [Alder, 1982]. For example, if two different sets of

numerical color differences calculated by color-difference equations A and B have the same number of color pairs, an F-test between these two color-difference equations can be easily performed using each equation's STRESS values:

$$F_{value} = \frac{STRESS_A^2}{STRESS_B^2} \tag{3.5}$$

where $STRESS_A$ and $STRESS_B$ are the STRESS values for measuring the deviation between color-difference equations A and B and visual color differences. The hypothesis that color-difference equations A and B are not significantly different will be rejected if $F_{value} < F_C$ or $F_{value} > 1/F_C$, where F_C is the critical values of the F distribution with certain confidence level and (j-1, j-1) degrees of freedom, where j is the total number of color difference pairs. F_C values can be calculated using the Matlab function of *finv((1-0.95)/2, j-1, j-1)*, if the confidence level is 95%.

Therefore, STRESS was used to measure the deviation between visual color differences and different kinds of numerical color differences: the T50-based ellipsoids, the 5000 randomized ellipsoids and three different color-difference equations.

3.6.1. Deviation between the RIT-DuPont Visual Color-Difference Data and Numerical Color-Difference Data

Every fitted ellipsoid (T50-based or randomized) equation was considered as one local color-difference equation, so the numerical color differences were calculated using the ellipsoid equations for the color-difference vectors of the RIT-DuPont dataset. The STRESS values measuring the deviation between the RIT-DuPont visual color differences (constant for all the color-difference vectors) and the numerical color differences calculated by each equation were calculated and are shown in Table 3.3.

Table 3.3. STRESS values between the RIT-DuPont visual color differences and the numerical color differences.

Color Center	STRESS values measuring the deviation between the RIT-DuPont visual data and …							
	T50-based Ellipsoids	5000 Randomized Ellipsoids				Color-Difference Equations		
		Mean	Max	Min	Std	ΔE^*_{ab}	ΔE^*_{94}	ΔE_{00}
1	7.75	8.74	22.38	7.74	0.88	44.25	31.88	13.43
2	3.13	5.06	11.33	3.16	1.28	23.53	12.83	13.02
3	4.01	4.68	6.96	4.02	0.43	15.18	14.39	16.97
4	3.30	4.67	15.24	3.31	1.45	27.63	8.77	10.34
5	3.56	4.63	10.82	3.59	0.81	19.58	6.68	10.73
6	3.09	3.73	6.95	3.10	0.45	36.24	20.46	20.82
7	6.16	7.39	13.05	6.16	1.05	21.35	9.41	8.14
8	5.55	5.82	6.88	5.54	0.19	20.19	12.99	9.23
9	2.71	6.72	18.06	2.82	2.27	20.92	9.99	13.77
10	1.49	3.05	7.05	1.54	0.77	14.12	14.32	13.60
11	0.00	3.85	11.38	0.54	1.57	19.40	11.69	12.01
12	5.02	5.80	9.10	5.03	0.51	16.94	16.43	14.70
13	0.00	5.39	17.26	0.37	2.76	44.14	21.92	26.19
14	7.76	13.61	35.73	7.77	4.92	24.54	23.10	20.40
15	1.07	8.04	26.78	1.16	5.23	15.46	11.54	12.86
16	0.99	4.98	16.39	1.07	2.47	17.62	9.41	6.44
17	1.18	5.97	16.35	1.43	2.44	22.66	23.30	4.75
18	0.84	4.68	11.61	0.97	1.66	19.47	19.30	20.73
19	1.87	8.59	23.15	2.05	4.31	35.87	18.28	8.27

For each color center, the STRESS value calculation included the color differences of all its color-difference vectors and the number of color-difference vectors included for each color center is shown in the last column of Table 3.4. Although it is known that ellipsoids are not the only possible shape describing equal color-difference contours [Stromeyer III, 1985], the small STRESS values provide evidence that, in general, the ellipsoids fit the visual color-difference vectors quite well. This is shown in the first column of Table 3.3 when STRESS was calculated for the T50-based ellipsoids. An ellipsoid seems an appropriate shape to represent contours of equal color difference.

Table 3.4. Results of the F-tests between three color-difference equations and each randomized ellipsoid.

Color Center	F-test results (%) of comparison between 5000 randomized ellipsoids and three color-difference equations (Deviation to the RIT-DuPont dataset)						95% Confidence Level		No. of the vectors
	ΔE^*_{ab}		ΔE^*_{94}		ΔE_{00}				
	Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*	
1	0.00	**100.00**	0.00	**99.94**	0.00	0.00	3.29	0.30	8
2	0.00	**100.00**	0.00	93.64	0.00	94.38	3.06	0.33	9
3	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	3.06	0.33	9
4	0.00	**100.00**	0.04	84.34	0.00	90.28	2.43	0.41	14
5	0.00	**100.00**	0.00	7.70	0.00	94.44	3.06	0.33	9
6	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	3.06	0.33	9
7	0.00	**99.86**	0.00	0.00	0.00	0.00	3.29	0.30	8
8	0.00	**100.00**	0.00	**100.00**	0.00	0.00	3.06	0.33	9
9	0.00	98.12	0.04	38.50	0.00	70.50	3.06	0.33	9
10	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	3.60	0.28	7
11	0.00	**99.84**	0.00	88.40	0.00	90.18	4.04	0.25	6
12	0.00	**99.96**	0.00	**99.94**	0.00	**99.50**	3.60	0.28	7
13	0.00	**100.00**	0.00	95.86	0.00	99.12	4.04	0.25	6
14	0.00	56.58	0.00	50.84	0.00	37.48	3.60	0.28	7
15	0.00	64.96	1.62	49.00	0.34	56.02	3.60	0.28	7
16	0.00	93.22	0.00	57.78	0.96	31.18	3.60	0.28	7
17	0.00	98.50	0.00	98.94	***12.32***	4.16	3.60	0.28	7
18	0.00	**99.80**	0.00	**99.76**	0.00	**99.92**	3.60	0.28	7
19	0.00	97.88	0.00	67.18	***9.22***	13.90	3.60	0.28	7

Although an ellipsoid might well fit the visual data, one cannot infer that the fitted ellipsoid defines the true visual contour, particularly if the visual data under-sample the local region. Color center 17, shown in Figure 3.3, is an example. Presumably, a vector sampling the principal axis would improve it reliability in representing the true visual contour. For color center 17, even though the T50 ellipsoid fit all the visual color-difference well, one cannot conclude that the ellipsoid defines the visual contour, as shown in Figure 3.3. Although the sampling scheme was designed to fit an ellipsoid and evaluate its size and orientation [Reniff, 1989], [Snyder, 1991], [Balonon-Rosen, 1993], this was not achieved for color center 17.

The STRESS values measuring the deviation between the visual and numerical color differences for the 5000 randomized ellipsoid equations were calculated for each

color center. The statistics of these STRESS values are also shown in Table 3.3. Most of the mean values of the STRESS values of randomized ellipsoids were close to the STRESS values of the T50-based ellipsoids, which verified the high precision of the RIT-DuPont dataset [Berns, 1991], [Berns, 1993], [Berns, 2002]. The maximum and minimum STRESS values of the randomized ellipsoids indicate the range of the possible equal color-difference ellipsoids if the visual experiment were repeated 5000 times. For color center 8, which had the tightest ellipse cloud in Figure 3.6, the mean, maximum and minimum STRESS values of the 5000 randomized ellipsoids were close to the STRESS values of the T50-based ellipsoids. For color centers with loose ellipse clouds in Figure 3.6, such as color centers 13, 14, 15 and 19, the mean STRESS values of the randomized ellipsoids were much larger than the STRESS values of the T50-based ellipsoids and the range between the maximum and minimum STRESS values were quite large compared with the STRESS values of the T50-based ellipsoids.

The STRESS values measuring the deviation between the visual and numerical color differences calculated by three CIE color-difference equations: ΔE^*_{ab}, ΔE^*_{94}, and ΔE_{00}, were also computed and shown in Table 3.3. Compared with the STRESS values of the T50-based ellipsoids, all the STRESS values of these three color-difference equations were quite large, which illustrated that the T50-based ellipsoids, as local color-difference equations, fitted the visual data much better than the overall color-difference equations. Since some of the STRESS values of CIE94 and CIEDE2000 were much smaller than the maximum STRESS values of the randomized ellipsoids, which represented the uncertainty limit of the RIT-DuPont data, there was the possibility that CIE94 and CIEDE2000 over-fitted the visual data at some color centers. A quantitative

method was developed, described below, that evaluated over- and under-fitting of the visual data.

F-tests between each of the three color-difference equations and each of the 5000 randomized ellipsoids were performed based on Equation (3.5) using the STRESS values calculated previously. An α of 0.05 was used in determining the critical values, $1/Fc$ and Fc, shown in Table 3.4 for each color center. There was a range of critical values because the number of vectors varied among the color centers. The percentage of the randomized ellipsoids whose STRESS values were significantly larger or smaller than the STRESS values of the color-difference equations were computed and shown in Table 3.4. For example, consider color center 1. 100% of the F-tests evaluating ΔE^*_{ab} were significantly larger than the critical value. A sphere (CIELAB ΔE^*_{ab}) was statistically inferior to a fitted ellipsoid in all cases. Since the visual data were generated randomly based on visual uncertainty, we assigned statistical significance to this percentage. If the percentage was greater than the critical value 99.5%, approximately equal to 3σ, then the equation under-fitted the visual data. This implies an opportunity for improvement. For CIE94, the percentage was 99.94%, and it did not fit the visual data. For CIEDE2000, none of the cases was significantly inferior to the fitted ellipsoids. Thus, CIEDE2000 well fit the visual data for color center 1. Color center 17 exemplifies a third possibility. For CIEDE2000, 12.32% of the cases had F values smaller than the critical value. Using the same 3σ, since this percentage exceeded 0.5% then the equation over-fitted the visual data. Over-fitting implies excessive model complexity. It is known that CIE94 is a very simple equation and using a 3σ critical value resulted in one color center being over-fit. Accordingly, the critical value was reduced to 2σ, 5%. If the percentage of cases where

the F-test value was less than the critical value exceeded 5%, then the equation over-fitted the visual data. This occurred for color centers 17 and 19 for CIEDE2000. For an ideal equation, the significance percentages would not be exceeded and values near zero are achieved. For color center 7, CIE94 and CIEDE2000 achieved this aim. In Table 3.4, italic and bold numbers mean the corresponding color-difference equation over-fitted the visual data, based on the critical value of 5%. Bold numbers mean the color-difference equation did not fit the visual data well enough, based on the critical value of 99.5%. Similar labels are used in Table 3.6.

3.7. Methods for Other Datasets

The equation-performance evaluation described above required stated uncertainty for each color-difference pair and generating a large number of ellipsoids. The datasets used commonly for equation evaluation, e.g. the Luo-Rigg [Luo, 1986], Witt [Witt, 1999] and Leeds [Kim, 1997a], [Kim, 1997b] datasets, only include an average standard error. Furthermore, the sampling may not enable ellipsoids to be fit locally. Thus, we sought methods that facilitated the comprehensive evaluation for these situations, described above.

3.7.1. Non-Ellipsoid Method

This method was designed for datasets with uncertainty information for each color pair, but where the sampling did not enable ellipsoid fitting in order to derive a local color-difference contour. Since each color-difference vector in the RIT-DuPont dataset represented an equal visual color difference of 1.02 ΔE^*_{ab} for a near-neutral color, the uncertainty of the color-difference vectors could also be represented by the uncertainty of

this visual color difference. Specifically, the upper and lower 95% confidence limit of the visual color difference (approximately twice the standard error) was described by the following equations, respectively:

$$\text{Upper 95\% confidence limit } = 1.02 \times \frac{UFL - T50}{T50} \tag{3.6}$$

$$\text{Lower 95\% confidence limit } = 1.02 \times \frac{T50 - LFL}{T50} \tag{3.7}$$

where UFL and LFL are the upper fiducial limit and lower fiducial limit of the color-difference vector, respectively; T50 is the median tolerance of the color-difference vectors.

For each color-difference vector, 5000 randomized visual color differences were generated based on their uncertainties calculated by Equations (3.6) and (3.7). The randomization was performed with the method by which the randomized color-difference vectors were calculated above. Making use of the advantage of STRESS, which cannot only calculate the deviation between numerical color differences and visual color difference, or between one set of numerical color differences and another set of numerical color difference, but also the deviation between two sets of visual color differences, the deviation between each set of the randomized visual color differences and the T50 visual color differences of each color center was calculated and represented by STRESS values. In this calculation, the four color vectors with infinite fiducial limits were not included due to the same reason as in the above method. The statistics of these STRESS values are shown in Table 3.5.

Table 3.5. STRESS values between the randomized visual color differences and the T50 visual color differences.

Color Center	STRESS values						
	Between the randomized and the T50 visual color differences				Between the T50 visual color difference and color-difference equations		
	Mean	Max	Min	Std	ΔE^*_{ab}	ΔE^*_{94}	ΔE_{00}
1	7.11	40.13	0.89	5.13	22.41	16.86	10.64
2	5.29	20.69	0.83	2.46	18.05	9.62	10.12
3	3.02	8.79	0.61	1.02	12.46	12.02	13.64
4	4.87	17.82	0.96	2.47	18.67	9.76	10.10
5	4.17	12.49	0.76	1.70	14.56	6.35	10.00
6	2.61	7.24	0.61	0.85	20.64	9.13	10.77
7	4.93	22.15	0.52	2.43	17.08	7.06	6.82
8	2.13	4.50	0.63	0.57	12.51	10.38	7.98
9	8.26	33.99	0.89	4.06	16.32	10.22	13.75
10	2.87	6.20	0.29	0.88	13.94	14.47	7.87
11	4.17	15.37	0.52	1.90	22.16	14.13	18.01
12	3.14	8.23	0.46	1.03	15.72	13.17	10.72
13	10.10	69.22	0.72	8.69	22.25	14.89	13.52
14	14.56	48.65	1.54	7.01	22.66	22.51	20.06
15	13.20	60.39	0.84	11.75	17.10	11.97	15.10
16	5.66	21.66	0.75	3.13	13.98	6.31	5.60
17	7.73	38.03	0.84	4.58	26.99	22.29	5.11
18	5.30	15.58	0.83	1.93	18.64	18.71	16.42
19	14.10	59.84	1.13	11.41	17.25	16.74	9.53

The STRESS values describing the deviation between the three CIE color-difference equations and the T50 visual color differences of the RIT-DuPont dataset are also shown in this table for comparative purposes. Compared with Table 3.3, the statistics of these STRESS values in Table 3.5 presented very similar trends describing the uncertainty of the RIT-DuPont dataset for each color center and the relationship between the three color-difference equations and the RIT-DuPont visual color differences. The ellipsoid fitting process in the previous technique averaged the error of the color-difference vectors. Thus the standard deviations in Table 3.5 were larger than the standard deviations in Table 3.3, which demonstrated a larger range of the STRESS value clouds.

F-tests between the STRESS values of the three color-difference equations and the STRESS values of each of the 5000 sets of randomized visual color differences to the T50 visual color difference were performed based on Equation (3.5) and the calculated STRESS values. The same critical values in Table 3.4 were used for each color center because the same numbers of color pairs were used. The percentage of the randomized visual color difference sets whose STRESS values were significantly larger or smaller than the STRESS values of the color-difference equations were computed and shown in Table 3.6.

Table 3.6. Results of the F-tests between three color-difference equations and each randomized visual color difference sets.

Color Center	Percentage of F-Test Results for 5000 Randomized visual color difference comparing with Different Color-Difference Equations (Deviation to the T50 visual color differences)						95% Confidence Level		No. of the vectors
	ΔE^*_{ab}		ΔE^*_{94}		ΔE_{00}				
	Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*	
1	0.00	96.48	0.00	89.62	0.22	57.68	3.29	0.30	8
2	0.00	97.18	0.00	71.42	0.00	72.64	3.06	0.33	9
3	0.00	**99.84**	0.00	**99.50**	0.00	**99.92**	3.06	0.33	9
4	0.00	**99.84**	0.10	63.16	0.00	72.44	2.43	0.41	14
5	0.00	99.06	0.00	31.32	0.00	74.72	3.06	0.33	9
6	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	3.06	0.33	9
7	0.00	94.68	0.02	42.60	0.06	30.98	3.29	0.30	8
8	0.00	**100.00**	0.00	**100.00**	0.00	**99.94**	3.06	0.33	9
9	0.00	73.54	0.66	20.80	0.06	42.78	3.06	0.33	9
10	0.00	**99.82**	0.00	**99.90**	0.00	**99.58**	3.60	0.28	7
11	0.00	93.40	0.00	61.00	0.00	61.54	4.04	0.25	6
12	0.00	**99.86**	0.00	**99.82**	0.00	99.22	3.60	0.28	7
13	0.00	82.70	0.04	56.82	0.00	64.04	4.04	0.25	6
14	0.00	29.84	0.00	26.82	0.00	20.08	3.60	0.28	7
15	***6.26***	39.32	***15.54***	22.70	***11.08***	29.68	3.60	0.28	7
16	0.00	74.98	0.00	34.78	0.98	13.76	3.60	0.28	7
17	0.00	74.46	0.00	75.98	***18.48***	0.86	3.60	0.28	7
18	0.00	91.18	0.00	91.02	0.00	94.40	3.60	0.28	7
19	0.00	68.30	2.44	38.52	***23.90***	4.84	3.60	0.28	7

The data in Table 3.6 were highly correlated with Table 3.4 except for slight differences. 18.48% and 23.90% of the cases for color centers 17 and 19, respectively, had ΔE_{00} superior to the randomized visual color difference in fitting the T50 visual color

difference, which indicated over-fitting of ΔE_{00} in these two color centers, as indicated by Table 3.4. As mentioned above, the standard deviations of the STRESS values of the randomized visual color differences were larger than the STRESS values of the randomized ellipsoids generated in the previous method. This difference resulted in the difference between Table 3.6 and Table 3.4. The direction of the data change from Table 3.4 to Table 3.6 was determined by the position of the STRESS values of the color difference equations in the uncertainty cloud. For example, if the color-difference equation had large STRESS values out of the uncertainty cloud, which presented the larger deviation than the randomized visual color differences to the T50 visual color difference, the enlargement of the uncertainty cloud increased the percentages of "significantly smaller" and decreased the percentage of "significantly larger". This trend was observed for ΔE^*_{ab}. On the other hand, if the color-difference equation had the deviation to the visual color difference comparable to the uncertainty cloud, the enlargement of the uncertainty cloud will increase both the percentages of "significantly larger" and "significantly smaller" as shown by some color centers for ΔE_{00}. The difference between Table 3.6 and Table 3.4 did not compromise the merit of this technique. The ellipsoid fitting process induced new errors but on the other hand the process unbiased the errors caused by the visual uncertainty in the psychophysical experiment with the reasonable assumption that the ellipsoids are the proper shape of equal color-difference contours. There is trade-off between these two methods. The choice between them should be determined by the property of the datasets.

3.7.2 Average Standard Error Method

If the dataset only had the average uncertainty for all color pairs rather than the detailed uncertainty information for each color pair, the upper and lower 95% confidence limits are calculated using Equations (3.8) and (3.9) with the same standard error (*SE*) for all color pairs:

$$\text{Upper 95\% confidence limit } = \Delta V \times (1 + 2SE) \tag{3.8}$$

$$\text{Lower 95\% confidence limit } = \Delta V \times (1 - 2SE) \tag{3.9}$$

where ΔV is the visual color difference of each color-difference vector. The randomization of the visual color differences using the same standard error for all color pairs was performed based on the similar method in the non-ellipsoid method. The RIT-DuPont dataset was also used in testing this average standard error method. The statistics of the STRESS values of the randomized visual color differences were calculated and shown in Table 3.7. In this calculation, the four color vectors with infinite fiducial limits were not included in order to keep the calculation comparable to above two methods. Unavoidably, part of the uncertainty information for each color center was lost because of using the average standard error for all the color pairs, and the STRESS values for the different color centers were nearly identical. Using the same method as the above techniques, the F-tests were performed and the percentage data are shown in Table 3.8.

Table 3.7. STRESS values between the randomized visual color differences (using average standard error for all color centers) and the T50 visual color difference of the RIT-DuPont dataset.

Color Center	STRESS values Between the randomized and original visual color differences			
	Mean	Max	Min	Std
1	4.98	11.13	0.78	1.35
2	5.01	10.38	1.45	1.28
3	5.03	10.15	1.33	1.27
4	5.18	9.71	2.07	1.05
5	5.03	9.87	1.10	1.28
6	5.04	9.97	0.79	1.28
7	4.94	10.49	1.12	1.34
8	4.97	9.98	1.41	1.26
9	5.02	9.85	0.83	1.27
10	4.85	12.00	1.02	1.44
11	4.77	10.47	0.73	1.56
12	4.87	12.34	0.95	1.44
13	4.80	12.27	0.72	1.54
14	4.87	10.63	0.91	1.44
15	4.91	10.53	1.02	1.46
16	4.91	11.15	0.81	1.44
17	4.85	11.15	0.74	1.44
18	4.91	11.57	0.98	1.44
19	4.88	10.67	0.57	1.43

Table 3.8. Results of the F-tests between three color-difference equations and each randomized visual color difference sets (using average standard error for all color centers) based on STRESS.

Color Center	Percentage of F-Test Results for 5000 Randomized visual color difference comparing with Different Color-Difference Equations (Deviation to the T50 visual color differences)						95% Confidence Level		No. of the vectors
	ΔE^*_{ab}		ΔE^*_{94}		ΔE_{00}				
	Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*	
1	0.00	100.00	0.00	100.00	0.00	80.92	3.29	0.30	8
2	0.00	100.00	0.00	83.52	0.00	85.50	3.06	0.33	9
3	0.00	94.70	0.00	89.46	0.00	97.28	3.06	0.33	9
4	0.00	100.00	0.00	49.58	0.00	77.00	2.43	0.41	14
5	0.00	99.90	0.00	6.14	0.00	54.94	3.06	0.33	9
6	0.00	100.00	0.00	99.96	0.00	99.90	3.06	0.33	9
7	0.00	99.94	0.00	26.92	0.00	14.34	3.29	0.30	8
8	0.00	99.96	0.00	78.52	0.00	28.34	3.06	0.33	9
9	0.00	100.00	0.00	51.22	0.00	92.02	3.06	0.33	9
10	0.00	76.40	0.00	77.68	0.00	69.80	3.60	0.28	7
11	0.00	93.24	0.00	42.96	0.00	43.82	4.04	0.25	6
12	0.00	92.58	0.00	89.92	0.00	79.84	3.60	0.28	7
13	0.00	100.00	0.00	98.44	0.00	99.78	4.04	0.25	6
14	0.00	99.98	0.00	99.86	0.00	99.06	3.60	0.28	7
15	0.00	84.30	0.00	45.34	0.00	62.50	3.60	0.28	7
16	0.00	94.42	0.00	25.80	0.00	4.72	3.60	0.28	7
17	0.00	99.74	0.00	99.82	0.02	1.10	3.60	0.28	7
18	0.00	97.84	0.00	97.78	0.00	99.04	3.60	0.28	7
19	0.00	100.00	0.00	96.36	0.00	16.08	3.60	0.28	7

Comparing with Table 3.4 and Table 3.6, similar trends are seen in Table 3.8. CIEDE2000 and CIE94 achieved smaller percentages of "significantly larger " than ΔE^*_{ab} for most of the color centers. The under-fitting performance comparison between CIE94 and CIEDE2000 in Table 3.8 provided the same information as in Table 3.4 for all color centers except for four. For color center 1, 8, 12, 14, 16, 17 and 19, CIEDE2000 had better under-fitting performance; for color center 2, 4, 5, 9, 11, 13, 15, 18, CIE94 had better under-fitting performance. For color center 17, CIEDE2000 still had the worst over-fitting performance, as indicated by both of above two methods. But, because of the detailed information loss caused by using the average standard error for all visual color data, the percentages in Table 3.8 provided the least sensitive information in evaluating the color-difference equations compared with the results given in Tables 3.4 and 3.6. However, the average standard error method still provided enough under-fitting information. Unfortunately, cases of over-fitting information could not be determined.

3.8. Conclusions

The RIT-DuPont dataset was designed to enable ellipsoid estimation in each sampled position in CIELAB. Plotting ellipsoids facilitates an effective visualization tool for determining whether a color space well predicts suprathreshold judgments of color difference; in this case the fitted ellipsoids are spherical with similar radii length. They also can be used to observe covariance, particularly between lightness (L^*) and chromaticness (a^*b^*). Finally, a color-difference equation can be evaluated by comparing its inherent contour (e.g., for CIELAB a sphere while for CIEDE2000 an ellipsoid) with the visual-based ellipsoid. Depending on the number of colorimetric data points and corresponding uncertainties about each point, the fitted ellipsoids may have poor stability

in terms of their size and orientation when such variability is taken into account. Accordingly, 5000 randomized ellipsoids for each of 19 color centers comprising the RIT-DuPont dataset were generated based on both the tolerance median (T50) and visual uncertainty (fiducial limits). When plotted as two-space projections, they provided a qualitative description of the ellipsoid reliability, this reliability dependent on visual uncertainty. A range of variability was found for this dataset.

The ellipsoids were considered as local color-difference equations. A color-difference equation that well predicts the visual data should have equivalent performance to the uncertainty cloud formed by the randomized ellipsoids. The range of the uncertainty ellipsoid cloud was shown in the analysis of the shape and the orientation of the randomized ellipsoids. STRESS was used in measuring the deviation between visual color differences and numerical color differences calculated by the T50 ellipsoids, the randomized ellipsoids and three color-difference equations. The STRESS values provided the numerical range of the uncertainty cloud. The F-tests using STRESS provided a more stable technique to evaluate the performance of fitting between different color-difference equations and the visual color differences of the RIT-DuPont dataset. This technique provided both over-fitting and under-fitting information for evaluating the color-difference equations. For CIEDE2000, both events occurred.

In some datasets used to evaluate equation performance, the sampling does not facilitate ellipsoid fitting. In addition, there may not be uncertainty estimates for each color difference, only an average, expressed by a standard error. Two methods were developed for these cases and compared with the technique using randomized ellipsoids. When individual uncertainties were known, the non-ellipsoid method had good

correlation with predicting both under- and over-fitting conditions. Comparing with the method with ellipsoid fitting, the non-ellipsoid method provided higher sensitivity since there is no ellipsoid fitting process that induced additional error. But, the ellipsoid fitting process provided the tool to unbias the error caused by the visual uncertainty with the reasonable assumption that the ellipsoids are the proper shape of an equal color-difference contour. The choice between these two approaches will be determined by the property of the dataset requiring analysis. When only average standard error was known, the simplification reduced correlation, only enabling under-fitting to be determined. Since the current philosophy in equation development is reducing complexity with concomitant equivalent or superior performance to CIEDE2000, the developed methodology including its simplification provides an important tool in model performance.

Future research will use this methodology in evaluating datasets other than the RIT-DuPont dataset and determine whether this can be used as either an objective function or constraint in model development.

4. Color-Difference Formula Performance for Several Datasets of Small Color Differences Based on Visual Uncertainty

Note that this Chapter was submitted as an article to Color Research and Application, authored by S. Shen and R. S. Berns, respectively. It has been accepted with major revision. The requirement for major revision was a result of the earlier manuscript, described in Chapter 3, being yet to appear in print. This chapter is the first version.

4.1. Abstract

Visual uncertainty, while reported, is not used routinely when evaluating color-difference formula performance in comparison to visual data; rather, data are analyzed assuming no uncertainty; that is, repeating the experiment would result in the identical average results. Previously, Shen and Berns developed a method to determine whether a color-difference formula was well-fitting, under-fitting, or over-fitting visual data when visual uncertainty was considered (described in chapter 3). This method used the Monte-Carlo technique where reported average standard error was used to generate 5000 random sets of visual data. STRESS was used to compare the random data to the published average data and to the performance of CIELAB, CIE94, and CIEDE2000 color-difference formulas. Three datasets were evaluated: BFD-P, Leeds, and Witt. For the BFD-P data, incorporating visual uncertainty led to the same performance results as the average results, that CIEDE2000 was an improvement over CIE94, which was an improvement over CIELAB. For the Witt data, incorporating visual uncertainty led to the same performance results as the average results, that CIEDE2000 and CIE94 had equivalent performance, both an improvement over CIELAB. However, both formulas under-fitted the visual results; thus,

neither formula was optimal. For the Leeds dataset, the visual uncertainty analysis didn't support the improvement of CIEDE2000 over CIE94 that occurred when evaluating the average results. Both formulas well fit the visual data. When considering experimental design, having more observers and repetitions with fewer color pairs results in lower uncertainty than the converse. Finally, average standard error could be used to approximate visual uncertainty defined using STRESS.

4.2. Introduction

The performance testing of the CIEDE2000 color-difference formula was carried out using the following datasets: BFD-P [Luo, 1986], RIT-DuPont [Berns, 1991], Witt [Witt, 1999], and Leeds [Kim, 1997a]. An F-test was used to compare the variances of a color-difference formula with visual variances of each dataset to assess statistical significance [CIE Publ. 142-2001, 2001]. In addition to comparisons with other formulas, reduced versions of CIEDE2000 were also evaluated, for example, replacing a weighting function with unity. This verified that each weighting function was statistically significant. Following the promulgation of CIEDE2000, it was discovered that the RIT-DuPont dataset, unfortunately, was used incorrectly. This problem was corrected and the analyses repeated [Melgosa, 2008], though in this case, F-tests were performed on the metric, STRESS [Garcia, 2007]. Fortunately, the results were the same. However, these analyses did not account for visual uncertainty. This corresponds to the assumption that if the visual experiments were repeated, the identical visual results would be obtained.

In chapter 3, techniques were developed to evaluate color-difference formula performance incorporating visual uncertainty. Three methods were described in order to

deal with different types of reported visual uncertainty. The RIT-DuPont dataset was first used to evaluate three color-difference formulas: CIELAB, CIE94 [Berns, 1993] and CIEDE2000 [Luo, 2001]. The randomized equal color-difference ellipsoids for each of 19 color centers comprising the RIT-DuPont dataset were generated based on the tolerance median and visual uncertainty. The ellipsoid formulas were considered as local color-difference formulas. STRESS [Garcia, 2007] was applied quantifying the deviation between visual color differences and numerical color differences calculated by the T50-based ellipsoids, the randomized ellipsoids and the color-difference formulas. For each color center, one STRESS value cloud was created and composed of the STRESS values quantifying the deviation between the numerical color difference calculated by each randomized ellipsoid and visual color difference. F-tests were performed to compare the STRESS values of the color-difference formulas and the STRESS cloud. The percentage of randomized ellipsoids that were beyond the critical *F* values was used as a metric for evaluating whether a color-difference formula was under- or over-fitting the visual data. The non-ellipsoid, average standard error method was also developed for cases when the dataset does not enable ellipsoid fitting and where the uncertainty has been reported only as an average standard error. This technique only enabled under-fitting to be quantified, a result of reduced uncertainty knowledge. These methods were developed using a dataset of constant visual difference, ΔV, a result of using the method of constant stimuli. For datasets with a range of ΔV, there is an assumption that visual uncertainty is independent of the magnitude of ΔV. This assumption is met for visual data based on the gray-scale method.

In this chapter, the third technique, the non-ellipsoid, average standard error method, was applied for evaluating color-difference formulas with three other small color-difference datasets: BFD-P, Leeds and Witt. In addition, the results were used to determine opportunities for improved performance when deriving new formulas.

4.3. Datasets

During the development and performance testing of CIEDE2000, the BFD-P, RIT-DuPont, Witt, and Leeds datasets were combined by rescaling to a reference color difference and weight averaging based on the number of color-difference pairs resulting in the COM-Weighted dataset [Melgosa, 2004], [Melgosa, 2008]. After that, the COM-Weighted dataset was widely used in developing and evaluating the color-difference formulas for small color difference [Luo, 2001], [Cui, 2002], [Luo, 2006]. A summary of these datasets is shown in Table 4.1 [Luo, 1986], [Berns, 1991], [Witt, 1999], [Kim, 1997a], [Kim, 2001]. The three datasets analyzed in this research used the gray-scale method enabling the Shen and Berns method in chapter 3 to be used to incorporate visual uncertainty. It is clear that each dataset has a unique set of parametric factors and accordingly, combining data may lead to erroneous performance testing. More problematic would be formula development where it is likely that weighting functions could be incorporating parametric factors. Thus in this research, each dataset was evaluated independently. In the previous publication, the RIT-DuPont dataset was evaluated. In this publication, the remaining three datasets were analyzed.

Table 4.1. Summary of the small color-difference datasets.

Datasets	Numbers of color vectors or color pairs	Color difference range (ΔE^*_{ab})	Sample materials	Psychophysical methods	Average standard error	Reference white point
RIT-DuPont	156	0.78-4.41	Glossy paint	Paired comparison	5.50%	[94.81,100.00,107.30]
BFD-P	2776	0.04-18.21	Various materials but relative scales of individual sets adjusted using textile samples	Various methods but the results were adjusted using gray scale method	8.90%	
BFD-D	2028	0.04-16.08				[94.81,100.00,107.30]
BFD-M	548	0.05-18.21				[94.65,100.00,103.97]
BFD-C	200	0.07-3.90				[98.07,100.00,118.23]
Leeds	307	0.40-4.74	Glossy paint	Gray scale and paired comparison	10.00%	[94.81,100.00,107.30]
Witt	418	0.12-10.63	Glossy paint	Gray scale	10.00%	[94.81,100.00,107.30]

4.4. Single White Point

The BFD-P dataset is a combined dataset of many earlier experiments and a limited visual experiment to enable data pooling [Luo, 1986]. Some of the color pairs belong to different datasets resulting in overlapped regions in color space. These color pairs were combined into one color center group using cluster analysis, described below. The BFD-P dataset is composed of three sub-datasets with different reference white points: BFD-D, BFD-C and BFD-M. Thus, before clustering, the colorimetric values of the BFD-C and the BFD-M datasets were converted to the colorimetric values under D65 with the same color appearance using a chromatic adaptation transformation [Fairchild, 2005] shown in Equations (4.1) – (4.5).

$$\begin{bmatrix} X_{D65} \\ Y_{D65} \\ Z_{D65} \end{bmatrix} = M_{CAT02}^{-1} \cdot M_{vonKries} \cdot M_{CAT02} \cdot \begin{bmatrix} X_{Ori} \\ Y_{Ori} \\ Z_{Ori} \end{bmatrix} \tag{4.1}$$

$$M_{CAT02} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix} \tag{4.2}$$

$$M_{vonKries} = \begin{bmatrix} \frac{R_{W2}}{R_{W1}} & 0 & 0 \\ 0 & \frac{G_{W2}}{G_{W1}} & 0 \\ 0 & 0 & \frac{B_{W2}}{B_{W1}} \end{bmatrix} \tag{4.3}$$

$$\begin{bmatrix} R_{W1} \\ G_{W1} \\ B_{W1} \end{bmatrix} = M_{CAT02} \cdot \begin{bmatrix} X_{Wori} \\ Y_{Wori} \\ Z_{Wori} \end{bmatrix} \tag{4.4}$$

$$\begin{bmatrix} R_{W2} \\ G_{W2} \\ B_{W2} \end{bmatrix} = M_{CAT02} \cdot \begin{bmatrix} X_{WD65} \\ Y_{WD65} \\ Z_{WD65} \end{bmatrix} \tag{4.5}$$

where M_{CAT02} is the chromatic adaptation matrix adopted by the CIECAM02 color appearance model; X_{WD65}, Y_{WD65} and Z_{WD65} are the tristimulus values of illuminant D65; X_{Wori}, Y_{Wori} and Z_{Wori} are the tristimulus values of the white points of the BFD-C or the BFD-M sub-datasets; $M_{vonKries}$ is the adaptation matrix based on von Kries' hypothesis [von Kries, 1902].

4.5. Color-Difference Data Clustering

The uncertainty analysis was developed for color-difference pairs dispersed in a local region of color space (equivalent to a color center in the RIT-DuPont dataset), ideally such that a contour of equal color difference could be estimated. This requires a minimum of six pairs. Hierarchical cluster analysis [Jain, 1988] was applied to each dataset to separate its color pairs into different color center groups in CIELAB space. The clustering was performed based on the Euclidean distance between the average CIELAB

coordinates of each color pair in CIELAB, that is, ΔE^*_{ab}. Only clusters having six or more color pairs were retained for analysis.

2776 color pairs comprising the BFD-P dataset were clustered into 163 groups. Only 81 groups had six or more color pairs, reducing the number of color pairs to 2606 color pairs. The average colorimetric values of the members in each group were calculated and listed in Table 4.2. In addition, for each group, the color differences between each color pair and the average color center were computed and the statistics of these color differences are also shown in Table 4.2. In Table 4.2, the 81 groups were roughly separated as five categories by their chroma and hue angle to facilitate the analysis: neutral ($C^* \leq 5.0\ \Delta E^*_{ab}$), greenish ($C^* > 5.0\ \Delta E^*_{ab}$, $225° \geq h^* > 135°$), bluish ($C^* > 5.0\ \Delta E^*_{ab}$, $315° \geq h^* > 225°$), reddish ($C^* > 5.0\ \Delta E^*_{ab}$, $45° \geq h^* > -45°$) and yellowish ($C^* > 5.0\ \Delta E^*_{ab}$, $135° \geq h^* > 45°$). In each category, the groups were in ascending order of the average lightness. The maximum and the average color differences between each color pair and the average color center of its group were 15.9 and 1.2 ΔE^*_{ab}. Five groups close to the five color centers recommended by CIE [Robertson, 1978] (No. 3, 18, 36, 56 and 81) had more than 100 color pairs and the largest group had 174 color pairs. Bold letters were used to label these color centers in Table 4.2. The color differences between each two average color centers were calculated and compared with the maximum color difference between each member and the average color center of a color pair group to confirm the validity of the clustering. For each color pair group, if the color difference between the average color center of this group and other 80 color centers were larger than the maximum color difference between each member and the average color center of this group, the grouping may be problematic for this group. A detailed analysis on the

relationship between each member and the average color center was performed on groups No. 4, No. 32 and No. 63 for the above problem. It was found that even though the color difference between two average color centers was smaller than the color difference between the average color centers and the color pairs, these color pairs had a much smaller color difference to their original color centers than to the others. The validity of the grouping was confirmed.

Table 4.2. Color pair groups with six or more pairs of the BFD-P dataset (1 of 2).

Color areas	Color pair groups	No. of Pairs	Average color centers of color pair groups					Color differences between each color pair and the average color center (ΔE^*_{ab})			
			L^*	a^*	b^*	C^*_{ab}	h_{ab}	Mean	Max	Min	Std
Neutral (4 groups)	1	31	38.35	0.27	0.15	0.31	29.27	0.15	2.27	0.08	0.39
	2	68	50.34	-0.91	-0.58	1.08	-147.36	3.72	5.85	1.59	1.42
	3	**174**	**62.75**	**0.28**	**0.36**	**0.46**	**52.27**	**2.64**	**8.38**	**0.34**	**2.40**
	4	62	87.82	-1.41	2.35	2.74	120.95	3.79	15.89	1.78	3.07
Greenish (23 groups)	5	20	19.68	-17.46	1.68	17.54	174.52	0.00	0.00	0.00	0.00
	6	18	27.65	-29.04	-11.68	31.30	-158.09	1.15	3.31	0.64	0.98
	7	29	29.70	-17.42	-15.53	23.33	-138.28	0.00	0.00	0.00	0.00
	8	17	35.17	-43.74	3.76	43.90	175.09	0.77	3.26	0.39	0.95
	9	20	37.63	-30.92	-1.59	30.96	-177.06	0.93	2.98	0.35	0.95
	10	32	38.20	-26.79	22.53	35.00	139.94	0.69	4.16	0.37	1.03
	11	38	39.23	-22.69	-22.57	32.01	-135.16	0.81	5.58	0.07	1.52
	12	30	40.70	-44.56	36.63	57.68	140.58	0.00	0.00	0.00	0.00
	13	30	42.48	-36.26	28.01	45.82	142.32	0.00	0.00	0.00	0.00
	14	29	47.91	-37.46	-21.52	43.21	-150.12	0.00	0.00	0.00	0.00
	15	30	47.94	-43.04	-1.09	43.05	-178.55	1.69	5.38	0.88	1.24
	16	16	54.69	-29.44	16.50	33.75	150.73	2.15	5.16	0.75	1.31
	17	21	55.31	-49.86	27.00	56.70	151.56	0.68	4.42	0.22	1.19
	18	**169**	**56.34**	**-31.49**	**0.46**	**31.49**	**179.16**	**0.40**	**3.15**	**0.31**	**0.34**
	19	6	56.85	-42.50	-7.00	43.07	-170.65	2.03	4.33	0.89	1.52
	20	6	60.88	-38.15	17.37	41.92	155.53	1.88	2.78	0.89	0.67
	21	42	60.94	-45.59	42.18	62.11	137.22	1.01	3.57	0.44	0.96
	22	10	62.72	-25.09	-7.85	26.29	-162.64	0.00	0.00	0.00	0.00
	23	12	63.09	-30.30	-4.75	30.67	-171.10	2.07	5.65	1.15	1.67
	24	15	63.17	-20.92	-15.13	25.82	-144.12	0.00	0.00	0.00	0.00
	25	23	64.13	-11.96	-7.53	14.13	-147.80	2.38	7.90	1.09	2.02
	26	13	64.25	-22.49	3.77	22.80	170.48	0.60	2.43	0.26	0.69
	27	51	71.55	-18.08	-0.66	18.09	-177.89	0.51	3.58	0.40	0.46

Table 4.2. Color pair groups with six or more pairs of the BFD-P (continued, 2 of 2).

Color areas	Color pair groups	No. of Pairs	Average color centers of color pair groups					Color differences between each color pair and the average color center (ΔE^*_{ab})			
			L^*	a^*	b^*	C^*_{ab}	h_{ab}	Mean	Max	Min	Std
	28	10	14.00	-3.99	-5.53	6.83	-125.82	0.00	0.00	0.00	0.00
	29	26	22.46	17.76	-34.36	38.68	-62.67	1.42	5.34	0.37	1.51
	30	11	22.72	20.09	-46.69	50.83	-66.72	0.00	0.00	0.00	0.00
	31	16	23.04	-3.31	-6.29	7.11	-117.77	1.97	3.20	1.41	0.85
	32	26	24.57	-12.51	-14.25	18.96	-131.28	4.02	10.12	0.80	2.30
	33	38	29.69	8.36	-38.41	39.31	-77.72	1.64	6.02	1.02	1.41
	34	20	32.21	-4.23	-42.40	42.61	-95.70	0.00	0.00	0.00	0.00
	35	8	33.23	-14.70	-20.26	25.03	-125.97	1.73	3.50	1.12	0.91
	36	**121**	**35.68**	**5.42**	**-31.04**	**31.51**	**-80.09**	**0.33**	**5.11**	**0.24**	**0.46**
Bluish (20 groups)	37	15	37.44	-26.57	-29.04	39.36	-132.46	0.00	0.00	0.00	0.00
	38	51	41.86	2.56	-17.14	17.33	-81.51	1.60	2.73	0.89	0.70
	39	30	44.63	-30.44	-32.63	44.63	-133.01	0.00	0.00	0.00	0.00
	40	46	46.90	10.95	-14.67	18.31	-53.28	1.08	2.88	0.73	0.52
	41	50	47.25	-5.09	-24.66	25.18	-101.67	1.04	6.28	0.46	1.01
	42	30	51.37	5.60	-43.30	43.66	-82.63	0.00	0.00	0.00	0.00
	43	20	53.25	-0.69	-7.04	7.08	-95.56	0.00	0.00	0.00	0.00
	44	18	63.26	-3.06	-8.57	9.10	-109.63	0.62	2.86	0.35	0.65
	45	20	73.79	-14.10	-21.76	25.93	-122.94	0.00	0.00	0.00	0.00
	46	13	80.55	-13.38	-14.86	19.99	-132.00	0.00	0.00	0.00	0.00
	47	11	83.11	-4.16	-7.95	8.97	-117.65	0.00	0.00	0.00	0.00
	48	20	16.39	20.79	-11.99	24.00	-29.96	0.00	0.00	0.00	0.00
	49	12	22.38	12.47	7.38	14.50	30.63	0.00	0.00	0.00	0.00
	50	30	32.92	59.29	-0.07	59.30	-0.06	0.00	0.00	0.00	0.00
	51	20	34.15	54.16	43.68	69.58	38.89	0.00	0.00	0.00	0.00
	52	9	34.62	31.20	-10.42	32.89	-18.47	1.46	2.75	0.84	0.57
	53	29	34.87	27.25	-21.10	34.46	-37.76	0.78	3.50	0.31	1.05
	54	19	36.92	47.13	17.98	50.44	20.88	1.67	7.02	0.77	1.58
Reddish (16 groups)	55	22	39.33	52.89	48.03	71.44	42.24	0.00	0.00	0.00	0.00
	56	**153**	**44.76**	**37.82**	**23.22**	**44.39**	**31.55**	**0.41**	**2.76**	**0.13**	**0.30**
	57	30	45.44	64.63	46.82	79.80	35.92	0.00	0.00	0.00	0.00
	58	7	50.13	25.39	2.75	25.54	6.17	4.11	7.16	2.19	1.72
	59	43	54.79	51.47	36.49	63.09	35.34	1.00	6.56	0.32	1.36
	60	34	60.93	26.05	-14.39	29.77	-28.92	1.33	5.81	0.76	1.47
	61	12	62.38	23.17	18.42	29.60	38.48	0.00	0.00	0.00	0.00
	62	27	64.07	33.70	2.63	33.80	4.47	2.70	7.08	1.77	1.92
	63	23	65.93	14.10	4.62	14.84	18.16	3.70	9.74	2.05	2.17
	64	29	31.93	7.79	30.17	31.16	75.53	0.00	0.00	0.00	0.00
	65	69	52.42	42.02	53.08	67.70	51.64	3.65	7.60	1.30	1.07
	66	17	59.41	-18.58	57.71	60.63	107.85	1.18	6.51	0.57	1.54
	67	12	60.14	-25.30	38.22	45.83	123.50	0.33	1.95	0.18	0.51
	68	10	60.26	-34.01	36.27	49.72	133.16	0.00	0.00	0.00	0.00
	69	10	60.75	-7.83	17.19	18.89	114.48	0.78	2.67	0.28	0.92
	70	13	61.14	1.31	65.77	65.78	88.86	1.25	4.19	0.59	1.31
	71	15	61.61	20.47	62.94	66.18	71.98	0.77	2.59	0.35	0.88
Yellowish (18 groups)	72	18	61.87	-10.46	40.94	42.25	104.33	0.90	3.84	0.36	1.21
	73	17	62.03	10.77	29.84	31.73	70.16	1.04	3.11	0.34	1.14
	74	15	62.48	8.57	9.13	12.52	46.81	0.83	1.74	0.54	0.46
	75	9	62.60	1.88	20.95	21.04	84.87	0.00	0.00	0.00	0.00
	76	57	64.97	30.78	35.47	46.96	49.06	1.88	5.35	0.80	1.44
	77	30	70.05	-2.03	43.39	43.44	92.68	0.40	2.97	0.21	0.70
	78	14	71.28	-7.05	44.79	45.34	98.95	0.00	0.00	0.00	0.00
	79	57	75.91	-3.50	84.52	84.59	92.37	4.18	5.90	0.80	1.28
	80	20	83.55	2.93	102.10	102.15	88.36	0.00	0.00	0.00	0.00
	81	**112**	**86.86**	**-6.23**	**46.86**	**47.27**	**97.57**	**0.99**	**4.04**	**0.78**	**0.46**

A similar procedure was performed for the Leeds and the Witt datasets. The 307 color pairs in the Leeds dataset were separated into 21 groups and ten of them had six or more color pairs, summarized in Table 4.3. 257 color pairs were used for evaluation. In Table 4.3, groups No. 4, 7, 9 and 10 were close to the CIE recommended color centers. Group No. 8 was considered as the recommended blue color center, but its average color center had obvious deviation to the recommended CIELAB values. For the Witt dataset, 418 color pairs were grouped into six color centers. This was unexpected since Witt only sampled the five CIE color centers. However, two color pairs from Witt's green color center could not be clustered into any of the five color centers using the methodology described above. These two pairs were discarded resulting in 416 color pairs, shown in Table 4.4.

Table 4.3. Color pair groups with six or more pairs of the Leeds dataset.

Color areas	Color pair groups	No. of color pairs	Average color centers of color pair groups					Color differences between each color pair and the average color center (ΔE^*_{ab})			
			L^*	a^*	b^*	C^*_{ab}	h_{ab}	Mean	Max	Min	Std
Neutral	1	11	28.12	-0.16	0.57	0.60	105.79	0.64	1.99	0.15	0.53
	2	9	39.81	-0.21	0.34	0.40	121.27	0.38	0.80	0.11	0.24
	3	54	49.82	-0.35	0.53	0.63	123.79	0.79	3.74	0.20	0.80
	4	**19**	**63.57**	**-0.25**	**-0.03**	**0.25**	**-173.17**	**2.81**	**4.09**	**1.75**	**0.77**
	5	10	78.21	-0.35	-0.50	0.61	-125.01	0.53	1.79	0.19	0.49
	6	9	87.96	-0.48	-0.44	0.65	-137.76	0.92	2.06	0.48	0.52
Greenish	**7**	**14**	**55.00**	**-31.85**	**0.06**	**31.85**	**179.98**	**0.66**	**1.67**	**0.30**	**0.46**
Bluish	**8**	**64**	**36.32**	**-4.30**	**-28.02**	**28.35**	**-98.77**	**2.71**	**4.72**	**0.57**	**1.09**
Reddish	**9**	**52**	**44.38**	**36.72**	**22.01**	**42.81**	**30.96**	**1.26**	**3.06**	**0.52**	**0.56**
Yellowish	**10**	**15**	**85.91**	**-6.41**	**45.82**	**46.26**	**98.02**	**0.76**	**1.91**	**0.11**	**0.57**

Table 4.4. Color pair groups with six or more pairs of the Witt dataset.

Color areas	Color pair groups	No. of color pairs	Average color centers of color pair groups					Color differences between each color pair and the average color center (ΔE^*_{ab})			
			L^*	a^*	b^*	C^*	h	Mean	Max	Min	Std
Neutral	1	85	62.70	0.11	0.24	0.27	66.55	0.68	1.81	0.17	0.51
Greenish	2	78	56.22	-31.21	0.48	31.21	179.21	1.01	3.28	0.22	0.91
Bluish	3	85	35.57	5.25	-31.35	31.79	-80.54	0.89	2.87	0.15	0.86
Reddish	4	85	44.59	37.18	23.15	43.80	31.92	1.26	3.98	0.26	1.12
Yellowish	5	83	86.84	-6.96	46.60	47.12	98.55	1.72	5.92	0.34	1.61

4.6. Performance Evaluation of Color Difference Formulas

Fitting ellipsoids requires a single average position, accomplished by translation in the fitting space [Melgosa, 1997]. For example, when fitting the RIT-DuPont vector data, the average and maximum translation was 0.67 and 2.72 ΔE^*_{ab}, respectively.[9] As seen in Tables 4.2 – 4.4 in the last four columns, many of the color pairs would require considerable translation. Because these large translations may range beyond a local linearity assumption and the non-ellipsoid method produced similar results to the ellipsoid method, the non-ellipsoid, average standard error method was applied for the color-difference formula evaluation using these datasets. That is, 5000 randomized visual color differences were generated for each color pair. The randomization was performed with the Monte-Carlo method [Alder, 1981] assuming a normal distribution with $\mu = \Delta V$ and $\sigma = SE \cdot \Delta V$, where ΔV is the visual color difference of each color pair and *SE* is the average standard error of each dataset as shown in Table 4.1. For each color pair within a group, one visual color difference was selected randomly from the 5000 visual color differences. This was repeated for each color pair within a group. The selection was repeated 5000 times resulting in 5000 sets of visual differences for each color group.

STRESS (STandardized REsidual Sum of Squares), a multivariate statistical function [Garcia, 2007], was used as the measure of deviation. Comparisons can be made between visual and color-difference formula data, between two different color-difference formulas, or between two different visual datasets. The first and last comparisons were performed. For each color pair group, the STRESS values were calculated between the original visual color differences and the randomized visual color differences, resulting in 5000 STRESS values, the statistics listed in Tables 4.5 – 4.7 for the BFD-P, Leeds, and

Witt datasets, respectively. STRESS was also calculated between the original visual color differences and either CIELAB, CIE94, or CIEDE2000, also listed in Tables 4.5 – 4.7. The average STRESS values between the randomized and the original visual color differences were 8.3 for BFD-P, 9.3 for Leeds, and 9.8 for Witt, reflecting the average standard errors listed in Table 4.1. For the BFD-P dataset (Table 4.5), the STRESS values of all three color-difference formulas were larger than the average STRESS values of the randomized visual color difference with only one exception, group No. 35. More than half of them were larger than the maximum STRESS values of the randomized visual color differences. For the Witt dataset (Table 4.7), all the STRESS values of CIEDE2000 were larger than the maximum values of the randomized color differences. This indicates that better formula fitting is possible for the BFD-P and the Witt datasets, particularly for color centers where the color-difference formula STRESS exceeded the maximum STRESS. For the Leeds dataset (Table 4.6), the STRESS values between the visual color difference and numerical color difference calculated by CIE94 and CIEDE2000 were close to the average STRESS values between the randomized and original visual color differences. For CIE94 and CIEDE2000, STRESS values for most groups were smaller than the maximum STRESS values of the randomized color differences, indicating that both CIE94 and CIEDE2000 well fitted the Leeds dataset and further improvement is unlikely.

Table 4.5. STRESS values of the BFD-P dataset (1 of 2).

Color areas	Color pair groups	STRESS values						
		Between the randomized and original visual color differences				Between the original visual color difference and color-difference formulas		
		Ave	Max	Min	Std	ΔE^*_{ab}	ΔE^*_{94}	ΔE_{00}
Neutral	1	8.56	13.49	4.95	1.25	24.30	25.05	20.65
	2	8.67	12.59	5.25	1.02	25.77	26.50	19.07
	3	**8.79**	**11.27**	**6.54**	**0.68**	**32.30**	**34.71**	**26.31**
	4	8.51	14.57	4.43	1.44	33.93	35.22	15.88
Greenish	5	8.31	16.03	3.07	1.71	17.14	9.48	15.69
	6	8.22	16.69	3.12	1.84	23.97	15.72	13.76
	7	8.61	14.57	4.19	1.28	20.16	18.72	19.86
	8	8.41	14.26	3.55	1.59	36.69	28.42	27.92
	9	7.85	19.70	2.86	2.05	43.66	14.51	18.45
	10	8.54	13.71	4.28	1.35	22.23	21.02	18.85
	11	8.63	12.58	5.06	1.13	32.13	28.84	28.77
	12	8.52	14.60	4.34	1.39	40.32	32.98	31.93
	13	8.57	13.72	4.28	1.32	19.25	21.22	20.22
	14	8.56	13.92	4.53	1.24	20.66	24.33	23.83
	15	7.76	20.13	2.85	2.29	32.70	12.73	12.13
	16	8.15	16.70	2.67	1.94	21.47	14.48	14.50
	17	7.88	18.29	2.80	2.09	20.66	9.91	9.76
	18	**8.74**	**12.00**	**6.17**	**0.82**	**21.99**	**26.42**	**24.26**
	19	7.48	19.47	1.08	2.62	18.52	16.84	16.64
	20	6.52	22.11	0.92	2.65	17.70	13.28	10.06
	21	8.39	15.44	4.24	1.51	33.76	20.02	19.82
	22	8.18	16.14	2.61	2.05	22.16	12.50	13.55
	23	7.60	19.95	2.51	2.17	14.43	10.40	8.81
	24	8.40	14.84	3.48	1.64	11.79	20.20	18.19
	25	8.36	19.24	3.52	1.68	12.69	17.46	13.99
	26	8.14	17.95	3.21	1.96	14.87	17.64	15.14
	27	8.55	17.06	5.05	1.31	17.90	19.29	15.64
Bluish	28	8.20	17.43	1.84	1.96	15.07	11.06	14.42
	29	8.35	17.24	2.87	1.66	45.93	34.74	26.18
	30	8.11	16.56	3.02	1.99	49.25	44.29	28.51
	31	7.54	18.38	2.50	2.24	11.12	17.15	8.94
	32	8.40	18.74	4.46	1.56	17.27	18.35	20.88
	33	8.64	14.09	5.06	1.18	43.87	39.30	33.44
	34	8.40	15.92	3.60	1.61	43.49	39.00	28.33
	35	7.91	19.02	1.11	2.34	6.45	9.36	5.02
	36	**8.75**	**11.82**	**6.06**	**0.77**	**38.09**	**30.26**	**23.99**
	37	8.34	14.92	2.99	1.69	13.07	18.73	15.20
	38	8.70	12.76	5.04	1.03	27.04	23.49	12.79
	39	8.54	14.29	4.52	1.35	24.77	28.13	27.80
	40	8.47	15.49	4.30	1.49	22.15	22.43	19.47
	41	8.61	13.97	5.01	1.18	25.12	30.20	27.66
	42	8.58	13.86	4.34	1.32	31.07	29.91	25.38
	43	8.20	18.44	3.41	1.80	20.54	19.36	19.66
	44	8.36	15.84	3.50	1.68	24.67	21.35	14.10
	45	8.29	16.12	3.09	1.75	20.14	15.92	19.04
	46	8.30	14.98	2.31	1.78	11.38	21.07	21.05
	47	8.21	16.38	2.59	1.90	17.28	17.65	15.85

Table 4.5. STRESS values of the BFD-P dataset (continued, 2 of 2).

Color areas	Color pair groups	STRESS values						
		Between the randomized and original visual color differences				Between the original visual color difference and color-difference formulas		
		Ave	Max	Min	Std	ΔE^*_{ab}	ΔE^*_{94}	ΔE_{00}
Reddish	48	8.31	15.19	3.49	1.75	19.61	27.35	20.41
	49	8.26	15.53	2.68	1.81	21.08	18.37	14.17
	50	8.54	13.85	4.70	1.28	25.89	17.41	17.97
	51	8.34	16.01	3.46	1.71	43.24	16.93	23.04
	52	7.06	21.64	1.66	2.47	28.80	9.79	11.58
	53	8.30	16.29	4.21	1.58	28.96	22.54	20.69
	54	8.37	15.60	3.30	1.67	35.13	17.04	20.77
	55	8.47	14.67	3.94	1.52	32.52	25.17	25.47
	56	**8.79**	**11.49**	**6.29**	**0.70**	**26.37**	**31.65**	**28.52**
	57	8.56	14.38	4.20	1.32	32.49	17.87	19.28
	58	7.73	18.18	0.99	2.38	11.57	12.79	8.74
	59	8.68	12.88	5.01	1.13	22.42	21.70	19.22
	60	8.55	14.33	4.24	1.37	12.43	28.00	23.39
	61	8.33	16.00	2.84	1.84	22.31	10.41	11.15
	62	7.90	18.46	2.47	2.29	16.32	20.56	15.78
	63	8.46	15.72	4.55	1.53	14.62	13.67	9.10
Yellowish	64	8.56	15.72	4.44	1.34	18.06	18.69	21.88
	65	8.72	13.19	5.89	0.96	27.02	24.44	23.19
	66	8.06	17.23	2.96	2.02	36.58	20.95	15.46
	67	8.18	18.88	2.09	1.93	17.93	8.50	8.41
	68	8.16	17.39	2.52	1.97	20.96	14.88	15.06
	69	7.47	19.64	2.14	2.44	18.20	11.32	9.31
	70	7.95	17.84	1.84	2.06	48.41	31.05	22.96
	71	8.00	21.03	3.02	2.00	31.03	25.68	15.88
	72	8.42	15.36	3.36	1.59	32.11	18.52	15.74
	73	8.14	16.93	3.10	1.96	24.25	32.41	21.70
	74	7.57	19.69	2.68	2.13	9.99	15.46	11.31
	75	8.07	18.19	1.75	2.08	16.49	8.23	10.63
	76	8.64	14.58	5.37	1.15	35.63	25.74	20.45
	77	8.53	13.92	4.16	1.35	29.60	39.78	32.92
	78	8.03	18.35	3.03	2.01	24.70	19.62	19.49
	79	8.61	13.68	5.17	1.13	34.92	32.37	22.57
	80	8.34	16.16	3.09	1.64	37.31	29.19	20.33
	81	**8.76**	**12.19**	**6.14**	**0.81**	**31.14**	**49.14**	**35.91**
Average		8.28	16.02	3.63	1.64	25.30	21.94	18.94

Table 4.6. STRESS values of the Leeds dataset.

Color areas	Color pair groups	STRESS values						
		Between the randomized and original visual color differences				Between the original visual color difference and color-difference formulas		
		Ave	Max	Min	Std	ΔE^*_{ab}	ΔE^*_{94}	ΔE_{00}
Neutral	1	9.09	19.24	2.65	2.29	12.42	12.43	12.05
	2	8.85	20.00	1.67	2.58	6.81	6.81	7.09
	3	9.79	13.52	5.99	1.04	23.81	24.20	17.40
	4	9.46	18.64	3.80	1.72	16.42	16.42	14.75
	5	8.96	17.79	2.37	2.30	14.76	14.76	14.54
	6	8.91	19.12	2.54	2.41	17.87	17.88	16.96
Greenish	7	9.28	17.95	3.38	2.03	35.51	16.59	16.87
Bluish	8	9.80	14.03	6.78	0.95	31.48	22.07	15.14
Reddish	9	9.76	14.03	6.33	1.13	36.92	19.11	19.39
Yellowish	10	9.34	17.69	3.70	2.00	35.53	21.57	15.07
Average		9.32	17.20	3.92	1.84	23.15	17.18	14.93

Table 4.7. STRESS values of the Witt dataset.

Color areas	Color pair groups	STRESS values						
		Between the randomized and original visual color differences				Between the original visual color difference and color-difference formulas		
		Ave	Max	Min	Std	ΔE^*_{ab}	ΔE^*_{94}	ΔE_{00}
Neutral	1	9.82	13.38	7.21	0.91	23.30	23.01	26.77
Greenish	2	9.78	14.37	6.38	1.07	44.88	20.27	21.78
Bluish	3	9.78	14.64	6.88	1.02	46.42	31.37	18.73
Reddish	4	9.80	14.76	6.72	0.98	28.70	13.69	15.56
Yellowish	5	9.79	14.19	6.92	1.00	43.21	26.51	23.88
Average		9.79	14.27	6.82	0.99	37.30	22.97	21.34

Compared to the RIT-DuPont and the Witt datasets, the BFD-P and the Leeds datasets had much different numbers of color pairs in different groups. It was found that groups with more color pairs had larger average STRESS values for the randomized visual color differences and a smaller range of values than those with fewer color pairs. For example, in Table 4.5, the STRESS values of the five color centers with more than 100 color pairs had slightly larger averages, much smaller ranges between maximum and minimum, and much smaller standard deviation values than the other color groups. Accordingly, the range of the critical values in the F-test of these groups was smaller, too. That means it is more difficult to properly fit these groups (not under- or over-fitted). With the assumption that the equal color-difference contour is an ellipsoid and thus has only six degrees-of-freedom, over sampling the color center in the visual color-difference

experiment will not improve the precision of the ellipsoidal equal color-difference contours achieved. It is suggested by these results that better precision can be achieved by reducing the visual uncertainty in the visual color-difference experiment. That is, more observers with proper color-space sampling will achieve a more valuable color-difference dataset.

An F-test was performed comparing the STRESS values between the randomized visual color differences and the original visual color difference with the STRESS values between the numerical color differences and the original visual color difference to determine the performance of each color-difference formula accounting for visual uncertainty. The capability of making statistical inferences is one advantage of STRESS compared with other widely used metrics such as PF and PF/3 [Melgosa, 2008], [Garcia, 2007]. For two sets of numerical color differences calculated by different color-difference formulas A and B with the same color-difference dataset, an F-test between these two color-difference formulas can be performed by calculating the F value using the following formula:

$$F_{value} = \frac{STRESS_A^2}{STRESS_B^2} \tag{4.6}$$

where $STRESS_A$ and $STRESS_B$ are the STRESS values for color-difference formulas A and B. The hypothesis that color-difference formulas A and B are not significantly different will be rejected if $F_{value} < F_C$ or $F_{value} > 1/F_C$, where F_C is the critical values of the F distribution with certain confidence level and (j-1, j-1) degrees of freedom, where j is the total number of color-difference pairs in the color-difference dataset.

The STRESS values of the 5000 randomized visual color-difference sets can be considered a STRESS value cloud representing the deviation of visual color difference caused by visual uncertainty. The information of under-fitting and over-fitting information of a color-difference formula can be provided by comparing the STRESS values of the color-difference formula with this cloud. For a color center, if the STRESS value of one color-difference formula is significantly smaller than most of the values in the STRESS value cloud, the color-difference formula has a high possibility of over-fitting the color-difference dataset. On the contrary, if the STRESS value of the color-difference formula is significantly larger than most of the STRESS cloud values, the color-difference formula has a high possibility of under-fitting the color-difference dataset. That is, there exists the possibility of an improved formula for a color center. The percentage of the 5000 randomized visual color differences sets whose STRESS values were significantly larger or smaller than the STRESS values of the color-difference formulas was computed and listed in Tables 4.8 – 4.10 for the BFD-P, Leeds, and Witt datasets, respectively. In Tables 4.8 – 4.10, the percentages in the column labeled as 'Smaller' represent the STRESS value of the corresponding color-difference formula that is significantly smaller than the listed percentage of the STRESS values in the cloud. Similar to the previous research in chapter 3 using the RIT-DuPont dataset, a confidence interval of 95% was adopted in determining the critical values for each color center. The critical values are shown in the last two columns of Tables 4.8 – 4.10 for the different color centers of three datasets. The under-fitting information is quantified by the percentage of "significantly larger" and the over-fitting information is represented by the percentage of the "significantly smaller". The ideal fitting is achieved when both of the

percentages are zero, which represents the color-difference formula that achieves equivalent performance to the uncertainty cloud [Witt, 1987]. As discussed in the previous chapter, the average standard error method only enabled under-fitting information to be determined where 99.5% was used as the threshold of under-fitting. That is, if the percentage is greater than 99.5%, the formula has under-fitted the visual color-difference data in this color center. Since there are 81 color centers in the BFD-P dataset, the raw number and percentage of color centers within each color group that exceeded the threshold are listed in Table 4.11. The Leeds and Witt datasets are also summarized in similar fashion in Table 4.11.

Table 4.8. F-test results for the BFD-P dataset (1 of 2).

Color areas	Color pair groups	Percentage of F-test results for 5000 randomized visual color difference comparing with different color-difference formulas						Critical values at 95% confidence level	
		ΔE^*_{ab}		ΔE^*_{94}		ΔE_{00}			
		Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*
Neutral	1	0.00	100.00	0.00	100.00	0.00	100.00	2.07	0.48
	2	0.00	100.00	0.00	100.00	0.00	100.00	1.62	0.62
	3	**0.00**	**100.00**	**0.00**	**100.00**	**0.00**	**100.00**	**1.35**	**0.74**
	4	0.00	100.00	0.00	100.00	0.00	99.14	1.66	0.60
Greenish	5	0.00	91.52	0.02	7.60	0.00	82.48	2.53	0.40
	6	0.00	99.76	0.00	78.52	0.00	57.26	2.67	0.37
	7	0.00	99.98	0.00	99.80	0.00	99.98	2.13	0.47
	8	0.00	100.00	0.00	100.00	0.00	100.00	2.76	0.36
	9	0.00	100.00	0.00	76.54	0.00	95.24	2.53	0.40
	10	0.00	100.00	0.00	100.00	0.00	99.82	2.05	0.49
	11	0.00	100.00	0.00	100.00	0.00	100.00	1.92	0.52
	12	0.00	100.00	0.00	100.00	0.00	100.00	2.10	0.48
	13	0.00	99.90	0.00	100.00	0.00	100.00	2.10	0.48
	14	0.00	100.00	0.00	100.00	0.00	100.00	2.13	0.47
	15	0.00	100.00	0.10	71.80	0.18	66.10	2.10	0.48
	16	0.00	98.18	0.00	61.00	0.00	61.32	2.86	0.35
	17	0.00	98.34	0.28	23.94	0.34	22.30	2.46	0.41
	18	**0.00**	**100.00**	**0.00**	**100.00**	**0.00**	**100.00**	**1.35**	**0.74**
	19	0.00	44.98	0.00	35.42	0.00	34.04	7.15	0.14
	20	0.00	56.94	0.00	30.28	0.00	13.04	7.15	0.14
	21	0.00	100.00	0.00	99.94	0.00	99.94	1.86	0.54
	22	0.00	91.58	0.00	17.24	0.00	25.08	4.03	0.25
	23	0.00	57.92	0.04	17.14	0.22	5.92	3.47	0.29
	24	0.00	17.14	0.00	97.16	0.00	90.00	2.98	0.34
	25	0.00	50.04	0.00	95.46	0.00	69.28	2.36	0.42
	26	0.00	54.50	0.00	81.00	0.00	58.08	3.28	0.31
	27	0.00	99.92	0.00	99.98	0.00	98.46	1.75	0.57
Bluish	28	0.00	37.40	0.00	8.30	0.00	31.32	4.03	0.25
	29	0.00	100.00	0.00	100.00	0.00	100.00	2.23	0.45

Table 4.8. F-test results for the BFD-P dataset (continued, 2 of 2).

Color areas	Color pair groups	Percentage of F-test results for 5000 randomized visual color difference comparing with different color-difference formulas						Critical values at 95% confidence level	
		ΔE^*_{ab}		ΔE^*_{94}		ΔE_{00}			
		Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*
Bluish	30	0.00	100.00	0.00	100.00	0.00	99.68	3.72	0.27
	31	0.00	38.08	0.00	87.94	0.66	13.74	2.86	0.35
	32	0.00	96.94	0.00	98.56	0.00	99.58	2.23	0.45
	33	0.00	100.00	0.00	100.00	0.00	100.00	1.92	0.52
	34	0.00	100.00	0.00	100.00	0.00	100.00	2.53	0.40
	35	1.08	0.48	0.00	3.96	8.24	0.06	4.99	0.20
	36	**0.00**	**100.00**	**0.00**	**100.00**	**0.00**	**100.00**	**1.43**	**0.70**
	37	0.00	34.50	0.00	92.26	0.00	62.84	2.98	0.34
	38	0.00	100.00	0.00	100.00	0.00	82.74	1.75	0.57
	39	0.00	100.00	0.00	100.00	0.00	100.00	2.10	0.48
	40	0.00	100.00	0.00	100.00	0.00	99.92	1.81	0.55
	41	0.00	100.00	0.00	100.00	0.00	100.00	1.76	0.57
	42	0.00	100.00	0.00	100.00	0.00	100.00	2.10	0.48
	43	0.00	98.82	0.00	97.56	0.00	97.94	2.53	0.40
	44	0.00	99.92	0.00	99.48	0.00	59.12	2.67	0.37
	45	0.00	98.72	0.00	84.50	0.00	97.24	2.53	0.40
	46	0.00	12.64	0.00	96.46	0.00	96.46	3.28	0.31
	47	0.00	66.78	0.00	69.84	0.00	52.16	3.72	0.27
Reddish	48	0.00	97.84	0.00	100.00	0.00	98.80	2.53	0.40
	49	0.00	94.66	0.00	81.48	0.00	37.20	3.47	0.29
	50	0.00	100.00	0.00	99.32	0.00	99.74	2.10	0.48
	51	0.00	100.00	0.00	90.80	0.00	99.90	2.53	0.40
	52	0.00	98.40	0.04	15.02	0.00	28.66	4.43	0.23
	53	0.00	100.00	0.00	99.94	0.00	99.66	2.13	0.47
	54	0.00	100.00	0.00	90.20	0.00	99.02	2.60	0.39
	55	0.00	100.00	0.00	100.00	0.00	100.00	2.41	0.42
	56	**0.00**	**100.00**	**0.00**	**100.00**	**0.00**	**100.00**	**1.38**	**0.73**
	57	0.00	100.00	0.00	99.34	0.00	99.86	2.10	0.48
	58	0.00	9.92	0.00	15.28	0.00	2.82	5.82	0.17
	59	0.00	100.00	0.00	100.00	0.00	100.00	1.85	0.54
	60	0.00	59.22	0.00	100.00	0.00	100.00	2.00	0.50
	61	0.00	97.04	0.00	6.22	0.00	9.86	3.47	0.29
	62	0.00	90.52	0.00	99.00	0.00	88.00	2.19	0.46
	63	0.00	77.04	0.00	63.34	0.20	3.28	2.36	0.42
Yellowish	64	0.00	99.38	0.00	99.70	0.00	99.98	2.13	0.47
	65	0.00	100.00	0.00	100.00	0.00	100.00	1.62	0.62
	66	0.00	100.00	0.00	97.48	0.00	74.84	2.76	0.36
	67	0.00	78.66	0.10	1.78	0.10	1.74	3.47	0.29
	68	0.00	87.58	0.00	36.78	0.00	38.60	4.03	0.25
	69	0.00	77.90	0.00	23.36	0.12	9.76	4.03	0.25
	70	0.00	100.00	0.00	99.98	0.00	97.60	3.28	0.31
	71	0.00	99.98	0.00	99.58	0.00	75.58	2.98	0.34
	72	0.00	100.00	0.00	95.94	0.00	79.00	2.67	0.37
	73	0.00	99.70	0.00	100.00	0.00	98.56	2.76	0.36
	74	0.16	19.34	0.00	77.84	0.02	35.02	2.98	0.34
	75	0.00	47.56	0.02	1.22	0.00	6.80	4.43	0.23
	76	0.00	100.00	0.00	100.00	0.00	100.00	1.70	0.59
	77	0.00	100.00	0.00	100.00	0.00	100.00	2.10	0.48
	78	0.00	99.08	0.00	92.84	0.00	92.40	3.12	0.32
	79	0.00	100.00	0.00	100.00	0.00	100.00	1.70	0.59
	80	0.00	100.00	0.00	100.00	0.00	99.34	2.53	0.40
	81	**0.00**	**100.00**	**0.00**	**100.00**	**0.00**	**100.00**	**1.45**	**0.69**
Average		0.02	86.16	0.01	80.47	0.12	75.88		

Table 4.9. F-test results for the Leeds dataset.

Color areas	Color pair groups	Percentage of F-test results for 5000 randomized visual color difference comparing with different color-difference formulas						Critical values at 95% confidence level	
		ΔE^*_{ab}		ΔE^*_{94}		ΔE_{00}			
		Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*
Neutral	1	0.00	11.42	0.00	11.42	0.00	9.42	3.72	0.27
	2	2.98	0.16	2.98	0.16	2.20	0.24	4.43	0.23
	3	0.00	100.00	0.00	100.00	0.00	99.90	1.72	0.58
	4	**0.00**	**67.22**	**0.00**	**67.22**	**0.00**	**44.44**	**2.60**	**0.39**
	5	0.00	25.76	0.00	25.76	0.00	23.94	4.03	0.25
	6	0.00	45.80	0.00	45.84	0.00	38.44	4.43	0.23
Greenish	**7**	**0.00**	**100.00**	**0.00**	**54.58**	**0.00**	**57.86**	**3.12**	**0.32**
Bluish	**8**	**0.00**	**100.00**	**0.00**	**100.00**	**0.00**	**97.70**	**1.65**	**0.61**
Reddish	**9**	**0.00**	**100.00**	**0.00**	**100.00**	**0.00**	**100.00**	**1.74**	**0.57**
Yellowish	**10**	**0.00**	**100.00**	**0.00**	**93.56**	**0.00**	**40.24**	**2.98**	**0.34**
Average		0.30	65.04	0.30	59.85	0.22	51.22		

Table 4.10. F-test results for the Witt dataset.

Color areas	Color pair groups	Percentage of F-test results for 5000 randomized visual color difference comparing with different color-difference formulas						Critical values at 95% confidence level	
		ΔE^*_{ab}		ΔE^*_{94}		ΔE_{00}			
		Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*
Neutral	1	0.00	100.00	0.00	100.00	0.00	100.00	1.54	0.65
Greenish	2	0.00	100.00	0.00	100.00	0.00	100.00	1.57	0.64
Bluish	3	0.00	100.00	0.00	100.00	0.00	100.00	1.54	0.65
Reddish	4	0.00	100.00	0.00	89.90	0.00	99.46	1.54	0.65
Yellowish	5	0.00	100.00	0.00	100.00	0.00	100.00	1.55	0.65
Average		0.00	100.00	0.00	97.98	0.00	99.89		

Table 4.11. Summary of the number of under-fitting color-pair groups and their percentage compared with the total number of groups for the BFD-P dataset, and the totals for all three datasets.

	ΔE^*_{ab}	ΔE^*_{94}	ΔE_{00}
Neutral	4 (100.0%)	4 (100.0%)	3 (75.0%)
Greenish	13 (56.5%)	10 (43.5%)	9 (39.1%)
Bluish	11 (55.0%)	10 (50.0%)	10 (50.0%)
Reddish	8 (50.0%)	6 (37.5%)	7 (43.8%)
Yellowish	11 (61.1%)	10 (55.6%)	6 (33.33%)
BFD-P TOTAL	47 (58.0%)	40 (49.4%)	35 (43.2%)
LEEDS TOTAL	5 (50.0%)	3 (30.0%)	3 (30.0%)
WITT TOTAL	5 (100%)	4 (80.0%)	4 (80.0%)

For the BFD-P dataset, there was the expected improvement in average performance as the color-difference formula increased in complexity from CIELAB to CIE94 to CIEDE2000. As shown in Table 4.8, the average percentage decreased from

86.16 to 80.47 to 75.88. On the other hand, evaluating each color group revealed surprising results. CIEDE2000 was derived, in particular, to improve performance for neutrals and bluish colors compared with CIE94. For the neutrals, three out of four color centers were under-fitted. For blues, ten out of twenty color centers were under-fitted; this was the identical performance to CIE94. The combination of large visual uncertainty and possible over-sampling resulted in appreciable under-fitting. The greenish and yellowish regions had the greatest improvement. For the Leeds dataset, there was the expected improvement in average performance as shown in Tables 4.9. CIEDE2000 had large improvement compared to the other formulas for the yellowish color center. Both CIE94 and CIEDE2000 had large improvement compared to CIELAB for the greenish color center. Interestingly, the neutral region was well fit for all three formulas except for group 3. Group 3 had 54 color pairs while the other neutral groups ranged from 9 to 19 pairs. In this case, large visual uncertainty and under-sampling obscured improvement of CIEDE2000 in five of six neutral color centers. The Witt dataset (Table 4.10) was under-fitted for all three formulas except for the reddish color center where both CIE94 and CIEDE2000 fitted the visual data. The combination of large visual uncertainty and possible over-sampling resulted in appreciable under-fitting.

The number of color centers that were under-fitted is listed in Table 4.11. The BFD-P data showed marked improvement is formula performance from CIELAB to CIE94 to CIEDE2000 where the percentage of under-fitting reduced from 58.0% to 49.4% to 43.2%. Still, there is quite an opportunity for improvement. For the Leeds data, the percentage of under-fitting reduced from 50.0% for CIELAB to 30% for both CIE94 and CIEDE2000. Further improvement seems unlikely. For the Witt dataset, the

percentage of under-fitting was 100% for CIELAB and 80% for CIE94 and CIEDE2000. None of these formulas fit the visual data when uncertainty was considered. As a comparison, the STRESS values for each dataset were calculated, listed in Table 4.12 and F-tests performed comparing the average visual results with each color-difference formula, the results given in Table 4.12. In Table 4.12, only for the selected color pairs in the above analysis were considered and the values in bold are these beyond the range of critical values. This is the usual statistical analysis for formula performance evaluation. For each individual dataset, CIE94 and CIEDE2000 were a significant improvement compared with CIELAB, well correlating with the under-fitting percentages. For the BFD-P and Leeds datasets, CIEDE2000 was a significant improvement compared with CIE94. This result correlates for BFD-P but does not support the under-fitting percentages for the Leeds dataset. For the Witt dataset, CIE94 and CIEDE2000 were equivalent, correlating with the under-fitting percentages.

Table 4.12. STRESS values for each dataset and F-test values.

STRESS		**F-Test values**		
BFD-P ($1/F_C$ = 1.0798, F_C = 0.9261)		**CIELAB**	**CIE94**	**CIEDE2000**
CIELAB	42.64	1.0000	**0.6490**	**0.4920**
CIE94	34.35	**1.5409**	1.0000	**0.7582**
CIEDE2000	29.91	**2.0324**	**1.3189**	1.0000
Leeds ($1/F_C$ = 1.2783, F_C = 0.7823)		**CIELAB**	**CIE94**	**CIEDE2000**
CIELAB	40.14	1.0000	**0.6245**	**0.2358**
CIE94	31.72	**1.6014**	1.0000	**0.3775**
CIEDE2000	19.49	**4.2416**	**2.6488**	1.0000
Witt ($1/F_C$ = 1.2125, F_C = 0.8247)		**CIELAB**	**CIE94**	**CIEDE2000**
CIELAB	51.91	1.0000	**0.3905**	**0.3367**
CIE94	32.44	**2.5606**	1.0000	0.8621
CIEDE2000	30.12	**2.9702**	1.1600	1.0000

4.7. Conclusions

Visual uncertainty, while reported, is not used routinely when evaluating color-difference formula performance in comparison to visual data; rather, the average results are analyzed. A Monte-Carlo-based method, termed the non-ellipsoid average standard error method, was developed to enable visual uncertainty to be incorporated during analysis. This method was used to analyze the performance of CIELAB, CIE94, and CIEDE2000 in predicting the BFD-P, Leeds, and Witt datasets. For the BFD-P data, incorporating visual uncertainty led to the same performance results as the average results, that CIEDE2000 was an improvement over CIE94, which was an improvement over CIELAB. It was also found that there is still an opportunity for further improvement over CIEDE2000, particularly for bluish and neutral colors. This was unexpected since CIEDE2000 was derived to address the poor performance of these two color regions. For the Witt data, incorporating visual uncertainty led to the same performance results as the average results, that CIEDE2000 and CIE94 had equivalent performance, both an improvement over CIELAB. However, both formulas under-fitted the visual results; thus, neither formula was optimal. For the Leeds dataset, the visual uncertainty analysis did not support the improvement of CIEDE2000 over CIE94 that occurred when evaluating the average results. Both formulas well fit the visual data. When considering visual uncertainty, formulas that compensate for chroma of the color-difference pair will well model the Leeds dataset.

This methodology enabled the average standard error to predict STRESS as a first-order approximation. Standard errors of 8.9%, 10%, and 10% resulted in average

STRESS values of 8.3, 9.3, and 9.8, respectively. Thus, visual uncertainty can be defined using STRESS.

Perhaps the most important conclusion concerns the trade-offs between the number of color-difference pairs and the amount of visual uncertainty. If a color center is under-sampled and has large visual uncertainty, most formulas will well fit the visual data. This was demonstrated using the Leeds dataset. If a color center is over-sampled and has large visual uncertainty, most formulas will under-fit the visual data. This was demonstrated using the Witt dataset. For the datasets analyzed in this research, reducing visual uncertainty had a greater impact on performance. It seems that for a fixed total number of observations, having more observers and repetitions with fewer color pairs would be preferred than the converse. Kuehni has arrived at a similar conclusion [Kuehni, 2008]. The difficulty is determining, for a specific color center, the optimal number of samples and observations under the constraint of the physical sample properties, the psychophysical method, and the observers. Based on the definition of standard error, doubling the number of observers only reduces uncertainty by $\sqrt{2}$ (1.4). It would be worthwhile to study a single color center with many observers and systematically vary the number of samples, their colorimetric location relative to a local equal-color-difference contour, and the number of observers to discover the optimal conditions. Perhaps the techniques used in this and our previous research (in chapter 3) can be applied to recent data collected at North Carolina State University [Shamey, 2008]. Unfortunately, it is uncertain whether these optimal values can be generalized to different experimental conditions.

5. IPT Based Euclidean Color Space and Color-Difference Formula

Note that this Chapter was written as a first draft for an article to be published in Color Research and Application, authored by S. Shen and R. S. Berns.

5.1. Abstract

A Euclidean color space, IPT-EUC ($I^E P^E T^E$, Euclidean IPT space) was developed based on an optimized small color-difference formula $\Delta E_{IPT\text{-}OPT}$ using the IPT space. The optimized color-difference formula has similar chromatic modeling to the CIE94 color-difference formula and a transformation function for the lightness channel. With the recent developed Euclidean color space deriving technique, a Euclidean color space was developed and optimized. A rotation matrix was applied to not only define the $+P^E$ axis of the IPT-EUC as unique red, but also to create a standard equal color-difference contour in the middle of the IPT-EUC. The performance evaluation was performed that included STRESS, F-test, hue constancy and equal color-difference ellipsoids. In addition, a deviation evaluation considering the visual uncertainty was performed. It was shown in the evaluation that the IPT-EUC space has quite good performance to evaluate color quality of small color differences and is comparable to the best available small color-difference formula, CIEDE2000. The IPT-EUC is also capable of describing color as a uniform color space. The IPT-EUC is a potential candidate for a unique color model for both describing color and measuring color quality.

5.2. Introduction

Seeking one color model that can predict and describe perceptual color in "all" applications is a long-time goal for color scientists. This goal was separated into two research fields: color-difference formula and uniform color space modeling, and color appearance modeling. The former is to find color-difference formulas mapping the perceptual color difference in certain color space [Berns, 2000]. CIELAB [CIE Publ. 15, 2004], which was recommended in 1976, was popularly used as the fundamental uniform color space and many color-difference formulas were developed based on it, e.g., CMC(l:c), BFD(l:c), CIE94, and CIEDE2000 [Clarke, 1984], [Luo, 1987a], [Berns, 1993], [Luo, 2001]. The latter focused on predicting the perceptual color change under different viewing conditions [Fairchild, 2005]. More than ten color appearance models were developed in last 30 years, e.g., Hunt et al., Nayatani et al., RLAB, CIECAM97s and CIECAM02 [Hunt, 1991], [Hunt, 1994], [Nayatani, 1990], [Fairchild, 1993], [Fairchild, 1996], [CIE Publ. 131, 1998], [CIE Publ. 159, 2004]. Conventionally, these two fields were considered as totally different; different purposes, different applications and different types of data were used to fit different models. But, in 1996, Luo, et al. demonstrated the possibility of using the single model to predict both color appearance and color difference [Luo, 1996]. After that, efforts were put on either using color appearance models to evaluating color differences [Li, 2003] or deriving uniform color spaces from color appearance models [Li, 2002], [Luo, 2006], [Berns, 2007].

One of the uniform color space modeling techniques was to construct Euclidean color spaces based on color-difference formulas. Either analytical [Völz, 1998], [Völz, 1999-2000], [Thomsen, 2000], [Völz, 2006] or computational methods [Urban, 2007a],

[Urban, 2007b] were applied to map the non-linear, non-uniform color spaces to linear and uniform color spaces based on the clues provided by different color-difference formulas, which were fit by color-difference datasets. Generally speaking, if there is a color-difference formula, a Euclidean color space could be derived based on it using either analytical or computational methods. The analytical form of the uniform color space is desirable since it provides clear physical meaning and is easy to implement. Thus the color-difference formula with the forms that can be analytically integrated is required for the technique.

In this chapter, a color-difference formula based on the color appearance model, IPT [Ebner, 1998b] was derived. IPT was selected since it is simple and has improved hue constancy. Furthermore, IPT space is based on the multi-stage vision model, which will entitle the derived color-difference formula not only an empirical formula but a theoretical one. In addition, a Euclidean color space was developed based on the new derived color-difference formula with the analytical method mentioned above. Furthermore, the parameters of the Euclidean color space were optimized to achieve better agreement with the color-difference formula. The color-difference formula and the Euclidean color space were evaluated and compared with other color-difference formulas and uniform color spaces on difference aspects.

5.3. Experimental Datasets

A COM-weighted dataset was widely used in recent color-difference formula modeling and evaluation for small color difference [Berns, 1993], [Luo, 2001], [Luo, 2006], [Oleari, 2008], [Melgosa, 2004], [Melgosa, 2008]. The COM-weighted dataset was composed of

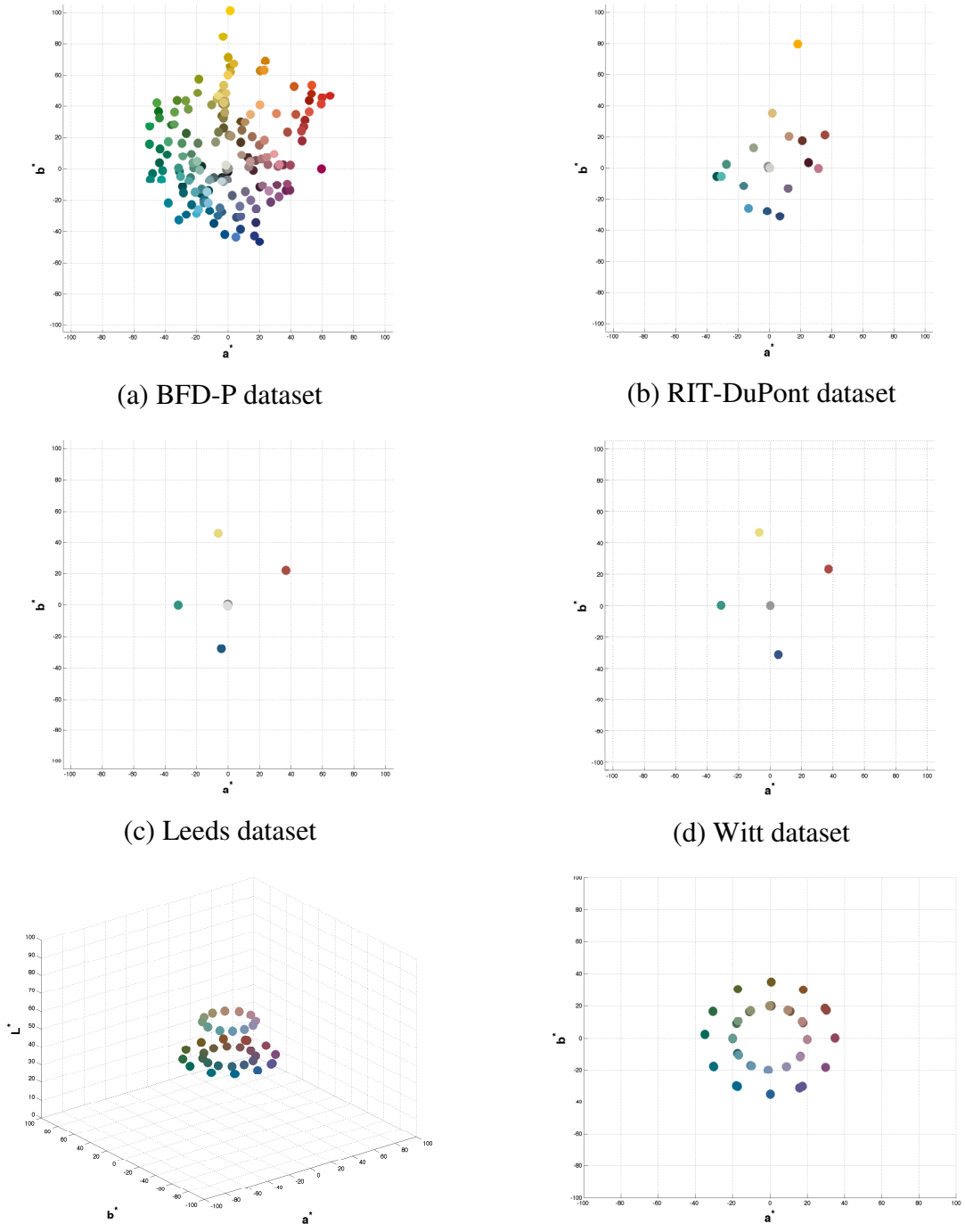
four separate datasets developed in four different laboratories: the RIT-DuPont [Berns, 1991], BFD-P [Luo, 1986], Leeds [Kim, 1997a] and Witt [Witt, 1999] datasets. Previously, these four datasets were used together by normalizing to a common anchor pair magnitude and were weighted based on their color pair numbers to balance their importance in the combined dataset. But, it was found later that there were inner differences between these four datasets [Berns, 1993], [Melgosa, 2004], [Melgosa, 2008] and they should be considered separately to avoid omitting important information. Thus, in this chapter, the BFD-P, Leeds and Witt datasets were used to fit the model separately. The RIT-DuPont dataset was combined with Qiao's hue suprathreshold color-difference dataset [Qiao, 1998] as another independent dataset (named as the RIT-DuPont-Qiao dataset in the following). The summary of these datasets is shown in Table 5.1.

Table 5.1. Summary of small color difference datasets.

Datasets	No. of pairs	Color difference range (ΔE_{ab}^*)	Pairs with $0.5 \le \Delta E_{ab}^* \le 5$			Sample materials	Psychophysical methods	Reference white point (*XYZ*)	Visual uncertainty
			% in all pairs	Range of lightness (CIELAB)	Range of chroma (CIELAB)				
RIT-DuPont-Qiao:	400*	0.78-4.41		13.09 - 84.70	0.14 - 86.13				
RIT-DuPont	312*	0.78-4.41	100.00%	13.09-84.70	0.14 - 86.13	Glossy paint	Pair comparison	[94.81,100.00,107.30]	5.5%
Qiao	88*	0.78-3.52		40, 60	20, 35	Glossy photo paper			11.5%
BFD-P:	2776	0.04-18.21	74.64%	0.90 - 92.76	0.06-104.62				
D65	2028	0.04-16.08	81.16%	0.90 - 92.76	0.06 - 85.23	Various materials and methods but relative scales of individual sets adjusted using textile samples and gray scale method.		[94.81,100.00,107.30]	8.9%
M	548	0.05-18.21	52.19%	26.88 - 77.27	0.23 - 92.14			[94.65,100.00,103.97]	
C	200	0.07-3.90	70.00%	15.22 - 87.68	3.53 - 104.62			[98.07,100.00,118.23]	
Leeds:	307	0.40-4.74	99.02%	27.35 - 90.22	0.19 - 50.25	Glossy paint	Gray scale and pair comparison	[94.81,100.00,107.30]	10%
Witt's:	418	0.12-10.63	86.60%	34.31 - 88.86	0.11 - 127.75	Glossy paint	Gray scale	[94.81,100.00,107.30]	10%

*: Both +T50 and –T50 are included.

The color centers of the RIT-DuPont, BFD-P, Leeds, Witt and Qiao datasets are shown in Figure 5.1.



(a) BFD-P dataset

(b) RIT-DuPont dataset

(c) Leeds dataset

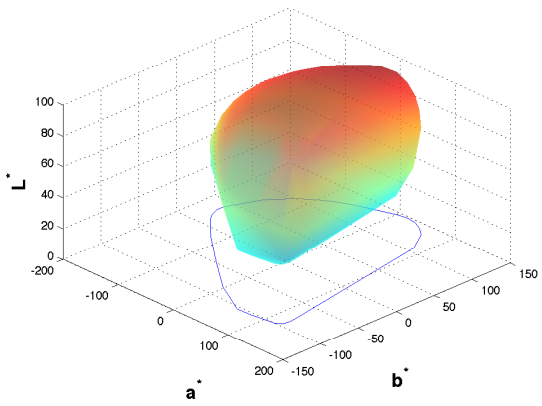
(d) Witt dataset

(e) Qiao dataset (3-D view)

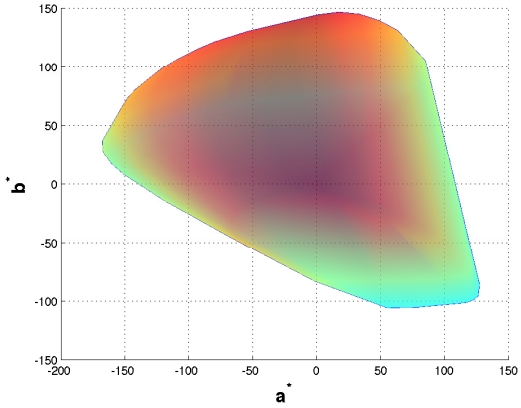
(f) Qiao dataset (a*-b* plane)

Figure 5.1. Color centers of small color-difference datasets. sRGB encoding.

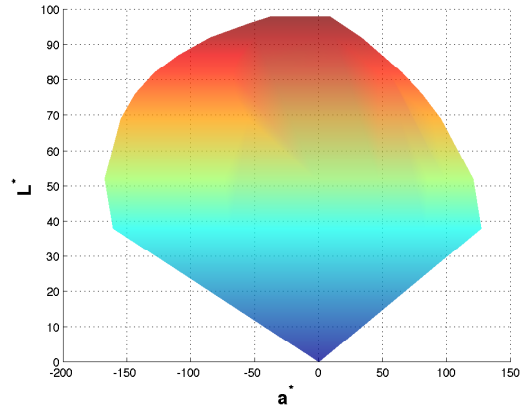
In Figure 5.1, the BFD-P, Leeds and Witt datasets were grouped first with clustering analysis described in Chapter 4 and only the average color centers of each color group were shown in the Figure 5.1. As shown in Figure 5.1, the BFD-P and RIT-DuPont datasets have the better sampling in CIELAB than the Leeds and Witt datasets. In addition, the Hung and Berns' [Hung, 1995], and Ebner and Fairchild's [Ebner, 1998a] constant hue datasets and OSA UCS sampling datasets [Moroney, 2003] were used in evaluating hue constancy performance of the uniform color spaces. Pointer's color gamut of real surface color [Pointer, 1980] was employed to evaluate the agreement between the derived color-difference formula and the derived uniform color space. Since the Pointer data only define the gamut larger than lightness of 38, the point of [0, 0, 0] in CIELAB was added to enlarge the gamut to the dark area. The final real surface color gamut is shown in Figure 5.2. It is shown in Figure 5.2 (b), (d) that boundary of bluish surface color is much smaller than other colors and yellowish surface color can achieve higher lightness.



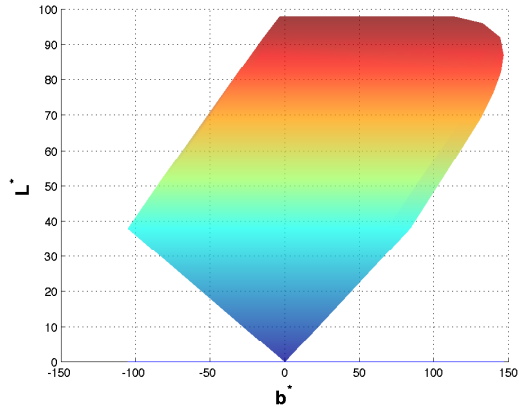
(a) Three-dimensional view



(b) a^*-b^* plane



(c) a^*-L^* plane



(d) b^*-L^* plane

Figure 5.2. Pointer's real surface color gamut.

5.4. Performance Evaluation Metric

The standardized residual sum of squares (STRESS), which is a multidimensional scaling technique, was used in evaluating the performance of color-difference formulas and uniform color spaces in this research [Garcia, 2007]. The definition of STRESS for measuring the deviation between visual color differences and numerical color differences is:

$$STRESS = 100 \cdot \left(\frac{\sum_i \left(\Delta E_i - F \Delta V_i \right)^2}{\sum_i F^2 \Delta V_i^2} \right)^{1/2}$$

$$\text{with } F = \frac{\sum_i \Delta E_i^2}{\sum_i \Delta E_i \Delta V_i} \tag{5.1}$$

where ΔE defines the numerical color difference, ΔV defines the visual color difference, F is a scaling factor for adjusting ΔV values to the same scale as ΔE, and i is the number of color-difference pairs.

STRESS has the advantage of making statistical inferences compared with other metrics that have been developed in the past such as PF/3, and others. For example, if two different sets of numerical color differences calculated by color-difference formulas A and B have the same number of color pairs, an F-test between these two color-difference formulas can be easily performed using each formula's STRESS values:

$$F_{value} = \frac{STRESS_A^2}{STRESS_B^2} \tag{5.2}$$

where $STRESS_A$ and $STRESS_B$ are the STRESS values for measuring the deviation between color-difference formulas A and B and visual color differences. The hypothesis that color-difference formulas A and B are not significantly different will be rejected if $F_{value} < F_C$ or $F_{value} > 1/F_C$, where F_C is the critical values of the F distribution with certain confidence level and (j-1, j-1) degrees of freedom, where j is the total number of color difference pairs.

In addition, the color-difference formula evaluating technique considering the uncertainty of the dataset introduced in Chapters 3 and 4 was performed on the color-difference formula and the Euclidean color space. In this technique, the uncertainties of the color-difference datasets were used to generate the STRESS value clouds that represented the capability of the visual color difference data being used in color-difference formula modeling. The STRESS values of different color-difference formulas were compared with these STRESS value clouds at different color centers to determine whether the color-difference formula over- or under-fit the visual color-difference data at different ranges of color space.

5.5. Color-Difference Formula

5.5.1. Formula Form and Color Space Modification

Equation (5.3) was finally decided as the form of the color-difference formula based on the IPT space. The color-difference formula was labeled as "$\Delta E_{IPT\text{-}OPT}$" standing for the "optimized color difference in IPT." This form has similar modeling to CIE94 on color difference along chroma and hue, which is simple and provides enough degrees of freedom to be fitted by the datasets. In addition, the formula can be analytically integrated, by which a Euclidean color space can be easily derived [Berns, 2007], [Thomsen, 2000]. A parametric factor k_L was applied on the lightness channel to compensate for differences of color materials.

$$\Delta E_{IPT-OPT} = \left\{ \left(\frac{\Delta I^E}{k_L} \right)^2 + \left(\frac{\Delta C_{P'T'}}{f_C\left(\overline{C_{P'T'}}\right)} \right)^2 + \left(\frac{\Delta H_{P'T'}}{f_H\left(\overline{C_{P'T'}}\right)} \right)^2 \right\}^{1/2} \tag{5.3}$$

where

$$C_{P'T'} = \sqrt{(P')^2 + (T')^2}$$

$$f_C(x) = 1 + \beta_C \cdot x$$

$$f_H(x) = 1 + \beta_H \cdot x$$

$$I^E = 100 \cdot \left[c \cdot \left(\frac{I'}{100} \right)^{a-1} \cdot \left(1 - \frac{I'}{100} \right)^{b-1} + \frac{I'}{100} \right] \tag{5.4}$$

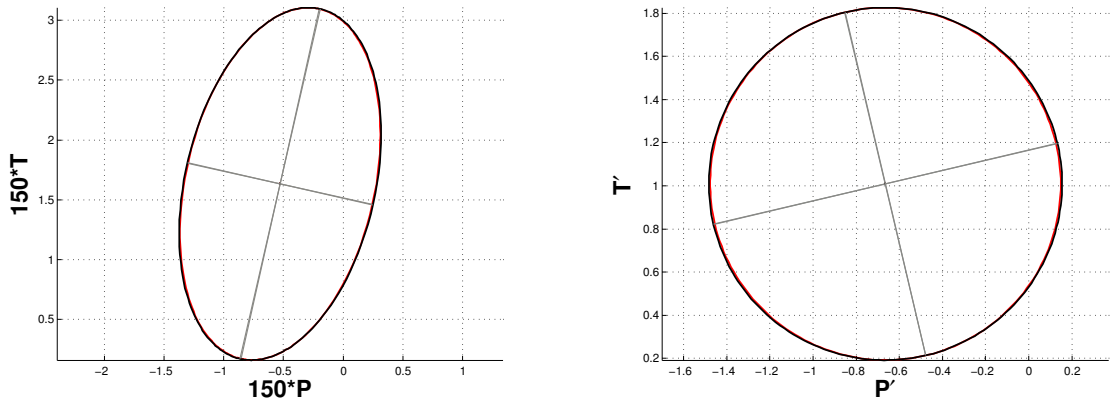
Before calculating the color difference with Equation (5.3), the IPT colorimetric attributes were first converted to the interim space *I'P'T'* with applying M_{COM}. The matrix M_{COM} was the multiplication of three matrices with different functions as shown in Equations (5.5) and (5.6).

$$\begin{bmatrix} I' \\ P' \\ T' \end{bmatrix} = M_{COM} \cdot \begin{bmatrix} I \\ P \\ T \end{bmatrix} \tag{5.5}$$

$$M_{COM} = M_{RED} * M_{Rot} * \begin{bmatrix} 100 & 0 & 0 \\ 0 & 150 & 0 \\ 0 & 0 & 150 \end{bmatrix} \tag{5.6}$$

The diagonal matrix at the right side of Equation (5.6) was to scale the range of IPT to roughly equivalent to CIELAB [Ebner, 1998b]. The purpose of the middle matrix M_{Rot} was to reshape the color space to achieve the equal numerical color differences in the medium gray along different hue directions in the chromatic plane. By doing this, a standard was achieved in the middle of the new color space and other equal color-difference ellipsoids can be compared with it. The coefficients of the M_{Rot} was calculated achieving both round projection and intersection on the chromatic planes of the fitted

equal color-difference ellipsoids for the color vectors around medium gray of the RIT-DuPont dataset (color center No. 3 in [Berns 1991]) in the IPT space. The fitted ellipses of medium gray color vectors of the RIT-DuPont dataset before and after applying M_{Rot} are shown in Figure 5.3.



(a) Before applying M_{Rot} (b) After applying M_{Rot}

Figure 5.3. Equal color-difference ellipsoid of the medium gray color center in the RIT-DuPont dataset.

Furthermore, the new color space was further rotated by applying the matrix M_{RED} to turn the hue angle of the Munsell Color 5R 14/4 (x = 0.5734, y = 0.3057 and Y = 12 under 1931 2 degree color matching functions and illuminant C, which was visually judged as best RED comparing with other red-hue patches in Munsell color system [Wyszecki, 2000] and Swedish NCS [Hård, 1996a], [Hård, 1996b]) to 0 degree. The coefficients of M_{Rot}, M_{RED} and M_{COM} are shown in Table 5.2.

Table 5.2. Matrices in $\Delta E_{IPT\text{-}OPT}$ color-difference formula.

M_{Rot}			M_{RED}			M_{COM} (after reducing significant figures)		
1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	100.0	0.0	0.0
0.0000	1.0000	−0.1474	0.0000	0.9721	0.2347	0.0	144.9	−2.1
0.0000	−0.0265	0.5507	0.0000	−0.2347	0.9721	0.0	−39.1	85.5

As shown in Table 5.2, the rotation and stretch resulted by employing M_{COM} operates only on the chromatic plane. A transformation curve shown in Equation (5.4) was applied onto the lightness channel to model the color-difference along lightness [von Seggern, 1990]. Comparing to the lightness transformation formulas introduced in previous research [Li, 2002], [Li, 2003], [Luo, 2006], [Cui, 2002], the formula of Equation (5.4) provided more degrees of freedom to describe the required conversion in the lightness channel.

5.5.2. Step-Wise Optimization of Color-Difference Formula

Besides the matrices, the coefficients (k_L, a, b, c, β_C and β_H) of the color-difference formula were optimized to minimize the STRESS values evaluating the deviation between the visual color difference and numerical color difference calculated using Equation (5.3) with four small color-difference datasets listed in Table 5.1. All visual color-difference datasets were transformed to IPT colorimetric attributes first. Since the transformation to IPT required that the input tristimulus values were under illuminant D65 and the 1931 2º observer, all color pairs in the four datasets were converted using the chromatic adaptation transformation shown in the Equations (5.7) – (5.11).

$$\begin{bmatrix} X_{D65} \\ Y_{D65} \\ Z_{D65} \end{bmatrix} = M_{CAT02}^{-1} \cdot M_{vonKries} \cdot M_{CAT02} \cdot \begin{bmatrix} X_{Ori} \\ Y_{Ori} \\ Z_{Ori} \end{bmatrix} \tag{5.7}$$

$$M_{CAT02} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix} \tag{5.8}$$

$$M_{vonKries} = \begin{bmatrix} \frac{R_{W2}}{R_{W1}} & 0 & 0 \\ 0 & \frac{G_{W2}}{G_{W1}} & 0 \\ 0 & 0 & \frac{B_{W2}}{B_{W1}} \end{bmatrix} \tag{5.9}$$

where

$$\begin{bmatrix} R_{W1} \\ G_{W1} \\ B_{W1} \end{bmatrix} = M_{CAT02} \cdot \begin{bmatrix} X_{Wori} \\ Y_{Wori} \\ Z_{Wori} \end{bmatrix} \tag{5.10}$$

$$\begin{bmatrix} R_{W2} \\ G_{W2} \\ B_{W2} \end{bmatrix} = M_{CAT02} \cdot \begin{bmatrix} X_{WD65} \\ Y_{WD65} \\ Z_{WD65} \end{bmatrix} \tag{5.11}$$

where M_{CAT02} is the chromatic adaptation matrix adopted by the CIECAM02 color appearance model [Fairchild, 2005]. X_{WD65}, Y_{WD65} and Z_{WD65} are the tristimulus values of illuminant D65 under 1931 2° observer: [95.047, 100, 108.883] X_{Wori}, Y_{Wori} and Z_{Wori} are the tristimulus values of the original white points. $M_{vonKries}$ is the adaptation matrix derived based on von Kries' hypothesis [von Kries, 1902]. In addition, only the color pairs with $0.5 \leq \Delta E_{ab}^{*} \leq 5$ in the datasets were employed in the modeling since the purpose was to develop color-difference formulas for small color-difference applications rather than large or threshold color-difference applications.

Six parameters in the formula were separated into three groups: chroma-related coefficients, β_C and β_H, lightness-related coefficients a, b, and c and parametric coefficient k_L. During the optimization, each group of coefficients was determined based on the optimization results of those datasets with enough distribution in the related

attributes. The BFD-P and RIT-DuPont-Qiao are most comprehensive datasets among these four, thus they were employed in the optimization for all parameters. Leeds and Witt datasets do not have enough distribution in the chroma plane and were not used in modeling chroma-related coefficients (β_C and β_H). Also, the lightness-related coefficients (a, b and c) are only decided by the BFD-P, RIT-DuPont-Qiao and Leeds since there are only five samplings along lightness in the Witt dataset. There are roughly two kinds of surface color materials in these four datasets: painting for the RIT-DuPont-Qiao, Leeds and Witt datasets; textile for the BFD-P dataset. The parametric coefficients were decided by these two groups of datasets for these two kinds of materials. To achieve better optimization results, the optimized parameters should be as few as possible. Thus a step-wise optimization was performed to achieve better results. First, all six coefficients in the formula were optimized for the BFD-P and RIT-DuPont-Qiao datasets and only chroma-related coefficients β_C and β_H were determined as the average of the results of these two datasets (shown in Table 5.3).

Table 5.3. First step optimization to determine the chroma-related coefficients.

	k_L	a	b	C	β_C	β_H	*STRES*
RIT-DuPont-Qiao	1.316	3.799	3.148	3.090	0.051	0.024	20.67
BFD-P	2.709	3.442	3.039	3.190	0.076	0037	27.72
				Average:	0.064	0.030	

Then, with these average coefficients β_C and β_H, all the other coefficients were optimized again for the BFD-P, RIT-DuPont-Qiao, Leeds and Witt datasets and the results are shown in Table 5.4.

Table 5.4. Second step optimization to determine the parametric coefficients.

	k_L	a	b	c	β_C	β_H	*STRES*
RIT-DuPont-Qiao	1.435	3.859	3.109	3.083	0.064	0.030	20.85
BFD-P	2.525	3.440	3.044	3.190			27.88
Leeds	2.006	4.320	2.881	2.981			22.42
Witt	1.507	4*	4*	3*			28.15
Parametric Coefficients:	1.6 for painting, 2.5 for textile						

*: The optimization results reached the predetermined boundaries.

The optimal lightness-related parameters for Witt dataset were much different than the other datasets since it does not have enough sampling of lightness. Thus a set of boundaries were set to limit the optimized coefficients into a range close to the results of the other datasets. Parametric coefficient k_L for textile surface color was determined by the BFD-P dataset and the k_L for painting surface color was computed from the average of the optimization results on the RIT-DuPont-Qiao, Leeds and Witt datasets. In the final step, only the lightness-related coefficients were optimized for all four datasets with previously calculated k_L, β_C and β_H and shown in Table 5.5.

Table 5.5. Third step optimization to determine the lightness-related coefficients.

	k_L	a	b	c	β_C	β_H	*STRESS*
RIT-DuPont-Qiao	1.6	3.901	2.852	3.119	0.064	0.030	21.03
BFD-P	2.5	3.415	3.042	3.204			27.88
Leeds	1.6	4	2.947	2.694			24.30
Averaged and adjusted:		4	3	3			
Witt*	1.6	4	3	3	0.064	0.030	31.88

*: The STRESS value of the Witt dataset was not optimized results. It is the calculating result based on the determined lightness transformation curve from optimization results of the other three datasets.

The Witt dataset doesn't have enough samples along lightness and it was decided the lightness-related coefficients were only determined by the results of the BFD-P, RIT-DuPont-Qiao and Leeds datasets. It was also found that the STRESS value of the Witt dataset was very sensitive to the change of these lightness-related coefficients. Thus, the average lightness-related coefficients optimized on these three datasets were slightly

adjusted to achieve acceptable fitting for the Witt dataset. The optimized lightness transformation curve is shown in Figure 5.4.

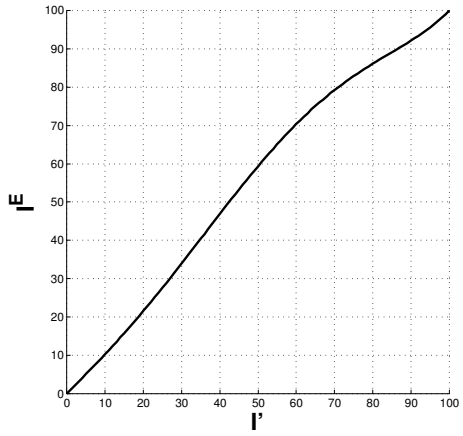


Figure 5.4. Optimized lightness transformation curve.

Finally, after canceling the insignificant digits, the coefficients for the small color-difference formula were shown in Table 5.6.

Table 5.6. Optimized coefficients for $\Delta E_{IPT\text{-}OPT}$ color-difference formula.

k_L	β_C	β_H	a	b	c
1.6 (painting) or 2.5 (textile)	0.06	0.03	4	3	3

5.6. Euclidean Color Space

5.6.1. Euclidean Color Space Derivation

The Euclidean color space was the ultimate goal of the research, by which the functions of both uniform color space and color-difference formula were realized. The Euclidean color space IPT-EUC (Euclidean IPT) was derived based on the color-difference formula derived above and shown in Equations (5.12) – (5.18).

$$\begin{bmatrix} I' \\ P' \\ T' \end{bmatrix} = M_{COM} \cdot \begin{bmatrix} I \\ P \\ T \end{bmatrix} \tag{5.12}$$

$$I^E = 100 \cdot \left[c \cdot \left(\frac{I'}{100} \right)^{a-1} \cdot \left(1 - \frac{I'}{100} \right)^{b-1} + \frac{I'}{100} \right] \tag{5.13}$$

$$C^E_{P'T'} = \left(\frac{1}{\beta_C} \right) \ln\left(1 + \beta_C \cdot C_{P'T'} \right) \tag{5.14}$$

$$h^E_{P'T'} = h_{P'T'} \tag{5.15}$$

$$P^E = P' \cdot \frac{C^E_{P'T'}}{C_{P'T'}} \tag{5.16}$$

$$T^E = T' \cdot \frac{C^E_{P'T'}}{C_{P'T'}} \tag{5.17}$$

$$\Delta E^E_{IPT-EUC} = \sqrt{\left(\frac{\Delta I^E}{k_L} \right)^2 + \left(\Delta P^E \right)^2 + \left(\Delta T^E \right)^2} \tag{5.18}$$

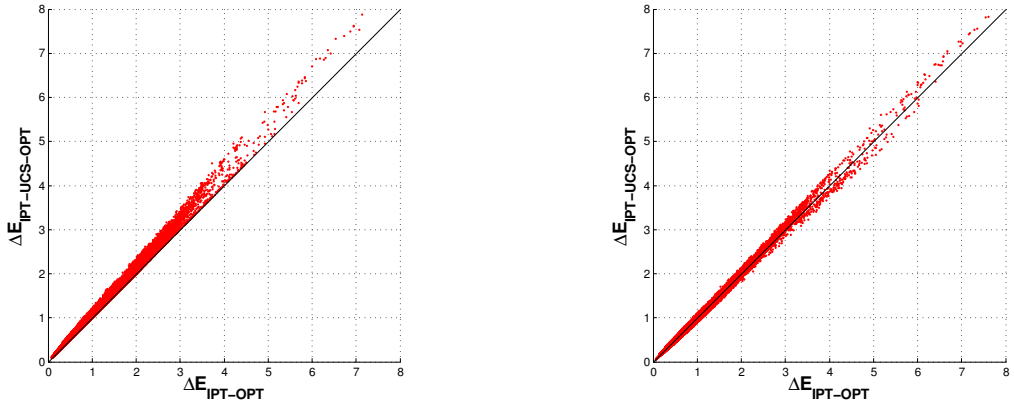
where IPT space was first modified by the rotation matrix M_{COM} in Equation (5.12) and the lightness transformation in Equation (5.13). The coefficients of them were the same as in $\Delta E_{IPT\text{-}OPT}$ color-difference formula shown in Table 5.2 and Table 5.6. Integration was performed on the chroma color-difference modeling shown in Equation (5.3) with the same fashion in previous research [Thomsen, 2000], [Berns, 2007] and shown in Equation (5.14). The new colorimetric attributes P^E, T^E were finally computed by applying the compression in Equations (5.16) and (5.17). The Euclidean distance $\Delta E^E_{IPT\text{-}EUC}$ was employed as the color quality metric.

5.6.2. Euclidean Color Space Optimization

The new Euclidean color space is not the exact transformation of the color-difference formula. Thus the numerical color differences calculated by Equation (5.18) were compared with the weighted color-difference formula in Equation (5.3) in similar fashion to Thomsen [Thomsen, 2000], shown in Figure 5.5 (a). A real surface color dataset was generated in CIELAB within the Pointer's real surface color gamut. First, CIELAB space was sampled evenly by 10,000 points. Then, Pointer's real surface color gamut was used as the boundary to select meaningful points from these 10,000 points. 1398 out of the 10,000 points within the gamut were selected. Each of these selected points was considered one color center, and 50 color vectors with random direction and random distance scaled between 0 and 5 ΔE^*_{ab} to the center were generated as 'batches'. STRESS value was employed again as the performance metric evaluating the deviation between the new Euclidean space and color-difference formula. The numerical color differences were converted to a wider range in the IPT-EUC space than in CIELAB. Thus, the maximum numerical color differences in Figure 5.5 (a) is around 8, larger than the maximum ΔE_{ab}^* of 5. Another optimization on the chroma compression parameter β_C was performed to minimize the deviation between the Euclidean distance in IPT-EUC ($\Delta E^E_{IPT\text{-}EUC}$) and the color differences calculation by $\Delta E_{IPT\text{-}OPT}$ with these color-difference pairs. By adjusting β_C from 0.06 to 0.071, the correlation between the Euclidean color space and the color-difference formula was improved and the new correlation is shown in Figure 5.5 (b). The final coefficients of the Euclidean color space are shown in Table 5.7.

Table 5.7. Optimized coefficients for the Euclidean color space IPT-EUC.

k_L	β_C	a	b	c
1.6 (painting) or 2.5 (textile)	0.071	4	3	3



(a) β_C=0.060 (b) β_C=0.071

Figure 5.5. Agreement between the Euclidean distance of IPT-EUC and the color differences calculated by $\Delta E_{IPT\text{-}OPT}$.

5.7. Performance of Euclidean Color Space and Color Difference Formula

5.7.1. STRESS value and F-test

The STRESS values evaluating the deviation between $\Delta E_{IPT\text{-}OPT}$ and $\Delta E^{E}_{IPT\text{-}EUC}$ were computed for the color vectors or pairs with 0.5<=ΔE_{ab}^{*}<= 5 in four small color-difference datasets. Besides the new Euclidean space and the color-difference formula, the STRESS values of CIELAB, CMC [Clarke, 1984], CIE94 [Berns, 1993], DIN99 [DIN6167, 2000], CIEDE2000 [Luo, 2001], DIN99d [Cui, 2002], CAM02-SCD, CAM02-UCS [Luo, 2006], OSA-GP [Huertas, 2006], CIECAM02-OPT series [Berns, 2007], OSA-UCS-E [Oleari, 2008], and DIN99o [Witt, 2009] color-difference formulas were also calculated. These STRESS values are shown in Table 5.8 by the order of their

introducing dates. For CIECAM02-based color-difference formulas, the viewing parameters for CIECAM02 attribute calculation are shown in Table 5.9 [Berns, 1991], [Luo, 1986], [Kim, 1997a], [Witt, 1999]. The calculation of OSA-UCS *LJG* attributes was based on the transformation formula proposed by Oleari in 2004 [Oleari, 2004]. In addition, before calculating the OSA-GP color difference, the chromatic adaptation of Equations (5.7) – (5.11) was performed for the BFD-C and BFD-M datasets to transform them to the colorimetric values under 1964 10° observer and D65. The best fit was achieved by CIEDE2000 for the BFD-P and Leeds dataset. For the RIT-DuPont-Qiao dataset, the smallest deviation was achieved with CIECAM02_OPT_ROT. DIN99d has the best performance for the Witt dataset.

Table 5.8. STRESS values of color-difference formulas for small color difference datasets with $0.5<=\Delta E_{ab}^{*}<=5$.

STRESS	**Datasets: $0.5<=\Delta E_{ab}^{*}<=5.0$**			
	BFD-P	**RIT-DuPont-Qiao**	**Leeds**	**Witt**
CIELAB	39.84	34.67	40.02	50.16
CMC	28.58	27.65	24.61	35.54
CIE94	33.22	25.10	30.49	31.85
DIN99	34.40	24.82	29.69	34.02
CIEDE2000	**27.51**	22.59	**19.11**	30.28
DIN99d	28.01	23.96	22.65	**29.92**
CAM02-SCD	28.39	26.39	22.00	30.49
CAM02-UCS	29.56	24.05	24.48	30.05
OSA-GP	28.88	25.12	27.30	32.26
CIECAM02_OPT	33.38	21.72	26.00	31.54
CIECAM02_OPT_ROT	36.19	**20.66**	29.56	34.55
CIECAM02_EUC	34.31	21.64	26.93	31.62
CIECAM02_EUC_ROT	37.22	20.96	30.55	34.84
OSA-UCS-E	27.63	25.71	26.10	35.35
DIN99o	27.74	26.31	23.57	30.81
$\Delta E_{IPT\text{-}OPT}$	28.19	21.10	24.39	32.24
$\Delta E^{E}_{IPT\text{-}EUC}$	28.09	21.32	24.23	31.11

Table 5.9. Viewing parameters for CIECAM02 calculations.

Datasets	Reference white point	L_A (cd/m^2)	Y_b	c	N_c
RIT-DuPont-Qiao:	[94.81, 100, 107.30]	127.32	10.89	0.69	1.0
BFD-P:					
D65	[94.81, 100, 107.30]	95.49	20.00	0.69	1.0
C	[98.07, 100, 118.23]	95.49	20.00	0.69	1.0
M	[94.65, 100, 103.97]	95.49	20.00	0.69	1.0
Leeds:	[94.81, 100, 107.30]	95.49	20.00	0.69	1.0
Witt's:	[94.81, 100, 107.30]	82.76	24.92	0.69	1.0

The STRESS values themselves are not enough to compare the performance of different color-difference formulas, thus the F-test was performed on the STRESS values in Table 5.8 and show in Table 5.10 – 5.13 to show whether the performance of these color-difference formulas are significantly different to each other (Table 5.10 – 5.18 are at the end of this chapter). In Table 5.10 – 5.13, the light shadow means the color-difference formula or uniform color space labeled on the left is significantly better than these labeled on the top. The dark shadow means the opposite. The $\Delta E_{IPT\text{-}OPT}$ and $\Delta E^{E}_{IPT\text{-}EUC}$ performed the same for all four datasets. In the F-test results for the BFD-P dataset shown in Table 5.10, the CMC, CIEDE2000, DIN99d, CAM02-SCD, OSA-UCS-E, DIN99o, $\Delta E_{IPT\text{-}OPT}$ and $\Delta E^{E}_{IPT\text{-}EUC}$ have similar performance and they are superior to the others. For the RIT-DuPont dataset in Table 5.11, $\Delta E_{IPT\text{-}OPT}$ and $\Delta E^{E}_{IPT\text{-}EUC}$ perform better than all the other color-difference formulas except for the CIEDE2000 and CIECAM02 series formulas. CIECAM02 series formulas developed by Berns and Xue are better than most of the other formulas for the RIT-DuPont-Qiao dataset but inferior in the evaluation based on other small color-difference datasets since these formulas were only optimized on the RIT-DuPont-Qiao dataset [Berns, 2007]. In Table 5.12, the CIEDE2000 is significantly better than all the other formulas for the Leeds dataset. The DIN99d, CAM02-SCD, CAM022-UCS, OSA-UCS-E, DIN99o, $\Delta E_{IPT\text{-}OPT}$ and $\Delta E^{E}_{IPT\text{-}EUC}$

have similar performance and better than others. Most of the formulas except for the CIELAB, CMC and CIECAM02 series formulas have similar performance for the Witt dataset shown in Table 5.13, which also indicated the difficulty to fit the Witt dataset. It was shown in the F-test that CIEDE2000 has the best performance for these four small color-difference datasets. The performance of the new developed $\Delta E_{IPT\text{-}OPT}$ and $\Delta E^{E}_{IPT\text{-}EUC}$ are superior or equal to all the other color-difference formulas listed except for the CIEDE2000 for the Leeds dataset. Especially, these two formulas performed very well for two most comprehensive datasets, the BFD-P and RIT-DuPont-Qiao datasets.

5.7.2. Performance Evaluation Considering Visual Uncertainty

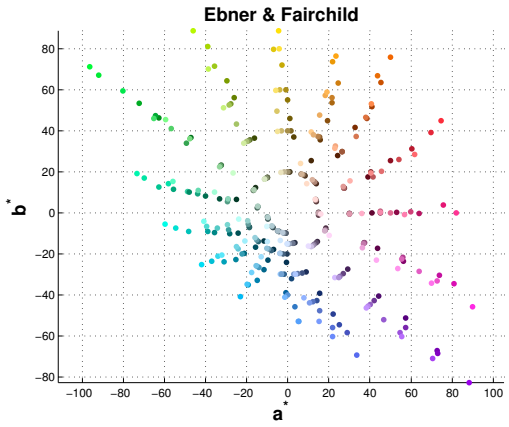
In the STRESS value and F-test evaluation, the uncertainty of the color-difference datasets was not considered. It is possible that some color-difference formulas over-fit certain dataset when their STRESS value is small. Thus, in addition to the F-test based on the STRESS values, the evaluating techniques introduced in Chapters 3 and 4, considering the visual uncertainty were performed on the $\Delta E_{IPT\text{-}OPT}$ and $\Delta E^{E}_{IPT\text{-}EUC}$ for the RIT-DuPont, BFD-P, Leeds and Witt datasets, shown in Table 5.14 – 5.17. The results of the similar evaluations performed in the Chapters 3 and 4 for the CIELAB, CIE94 and CIEDE2000 are also shown in these tables for comparison. Given the property of each dataset, the ellipsoidal method was used for the RIT-DuPont dataset and the non-ellipsoidal, average standard error method was applied for the other three datasets. In Tables 5.14 – 5.17, the same thresholds for judging over-fitting and under-fitting as in Chapters 3 and 4 were used. That is, if the number under the label of "smaller" is larger than 5 (which has the light shadow), the corresponding formulas over-fits the dataset at this color center. If the number under the label of "larger" is larger than 99.5 (which has

dark shadow), the corresponding formula under-fits the dataset at this color center. It was shown in the Chapter 3 that the visual data at the No. 17 and 19 color centers of the RIT-DuPont data were over-fit by the CIEDE2000. In Table 5.14, the over-fitting performance of the $\Delta E_{IPT\text{-}OPT}$ and $\Delta E^{E}_{IPT\text{-}EUC}$ are better than the CIEDE2000 on the color center No. 17 and 19. But the under-fitting performance is slightly worse than CIEDE2000 for the RIT-DuPont dataset. For the BFD-P, Leeds and Witt datasets, because of the property of the datasets, only the under-fitting information is achieved in the Table 5.14 – 5.17. For the BFD-P dataset, the $\Delta E_{IPT\text{-}OPT}$ and $\Delta E^{E}_{IPT\text{-}EUC}$ have slightly worse over-fitting performance than CIEDE2000 on the neutral and greenish color centers, but slightly better over-fitting performance on the reddish and bluish color centers. For yellowish color centers, the $\Delta E_{IPT\text{-}OPT}$, $\Delta E^{E}_{IPT\text{-}EUC}$ and CIEDE2000 have similar over-fitting performance. For the Leeds dataset, the over-fitting performance of CIEDE2000 is only slightly better than $\Delta E_{IPT\text{-}OPT}$ and $\Delta E^{E}_{IPT\text{-}EUC}$ on bluish color center, which reflected the fact that in deriving CIEDE2000, a great effort was put on modeling the Leeds dataset on bluish color centers. The $\Delta E_{IPT\text{-}OPT}$ and $\Delta E^{E}_{IPT\text{-}EUC}$ have similar under-fitting performance to the CIEDE2000 and CIE94 on the Witt dataset.

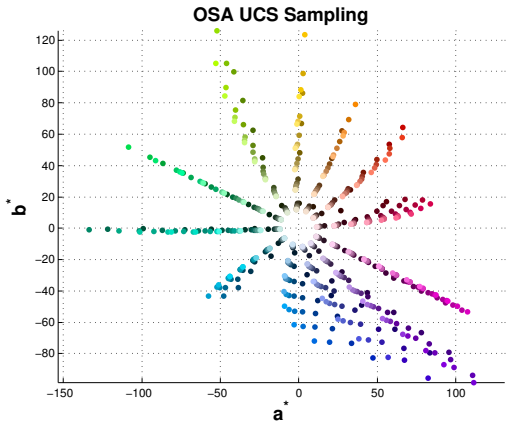
5.7.3. Hue Constancy

Three constant hue datasets, Hung-Berns, Ebner-Fairchild and OSA UCS datasets are plotted in the IPT-EUC space in Figure 5.6 to show the hue constancy of the Euclidean color space. In addition, as comparison, the datasets are also plotted in the CIELAB, IPT, CIECAM02 and OSA UCS color spaces in Figure 5.6. The IPT space in Figure 5.6 (c) has the best hue constancy property that is represented by the quite straight line along each hue angle. Only in purplish area (hue angle: 270° – 315°), a small amount of scatter

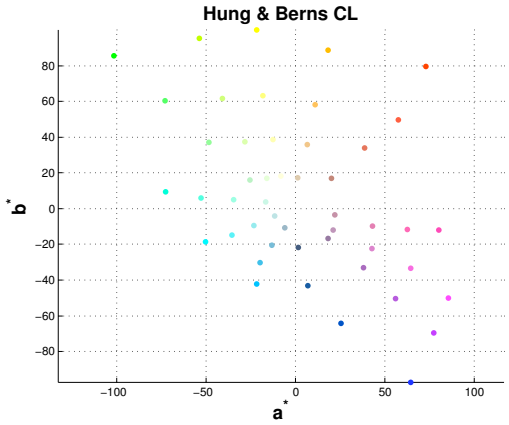
was observed. OSA-UCS has good hue constancy except for the greenish area (hue angle: 135° – 180°). In the greenish, bluish and purplish areas (hue angle: 180° – 315°) in the CIECAM02 *JCh* space, the hue constancy datasets scattered in certain range. The CIELAB space has the worst hue constancy performance comparing with others, especially in the bluish and purplish regions. The Euclidean color space IPT-EUC has good hue constancy performance in greenish, yellowish and reddish regions. But the datasets do not distribute as lines at bluish and purplish areas. Also, it was shown in the comparison between the circles formed by the least chroma data points in difference spaces that the directional stretch and rotation caused by the rotation matrix M_{COM}. The hue constancy performance of the IPT-EUC is worse than IPT space which was specially developed to keep the hue constancy, but much better the CIELAB spaces and comparable to that of CIECAM02 *JCh* and OSA-UCS spaces. Also, in the IPT-EUC space, the Munsell color 5R 4/14 is on the 0° hue angle axis which gave the $+P^E$ axis a physical meaning.



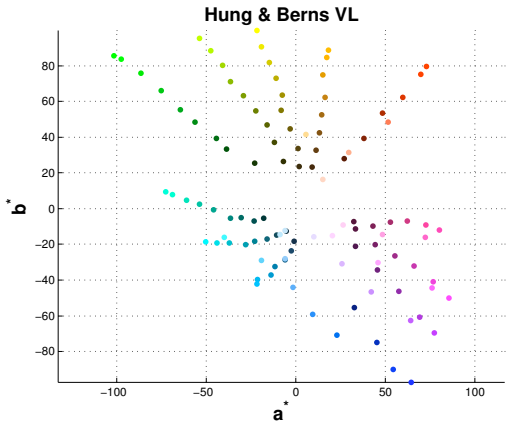
Ebner-Fairchild dataset



OSA UCS radial sampling



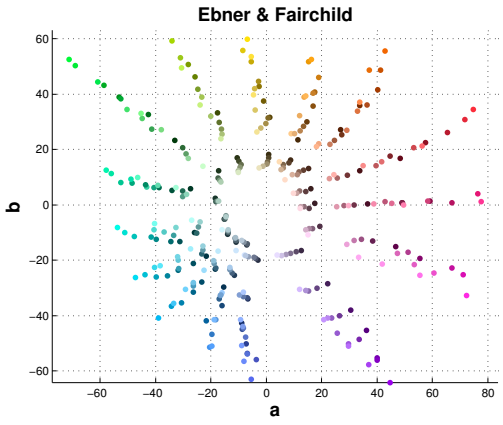
Hung-Berns constant lightness dataset



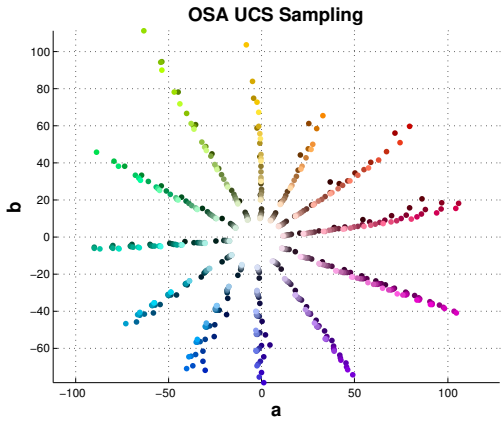
Hung-Berns varying lightness dataset

(a) CIELAB space

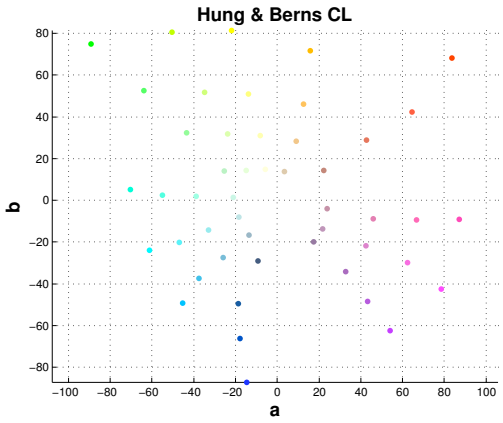
Figure 5.6. Constant hue datasets on different color spaces (part 1 of 5).



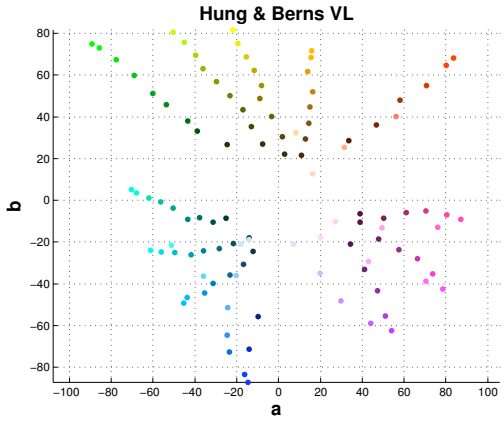
Ebner-Fairchild dataset



OSA UCS radial sampling



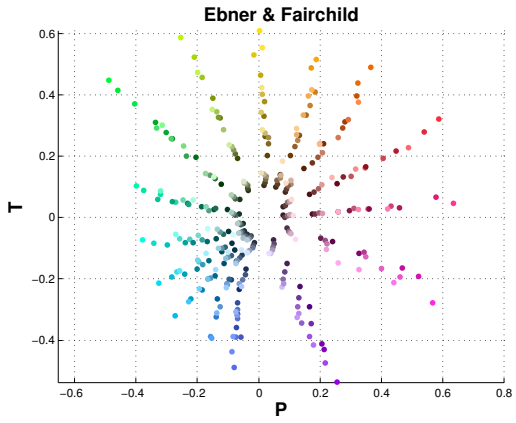
Hung-Berns constant lightness dataset



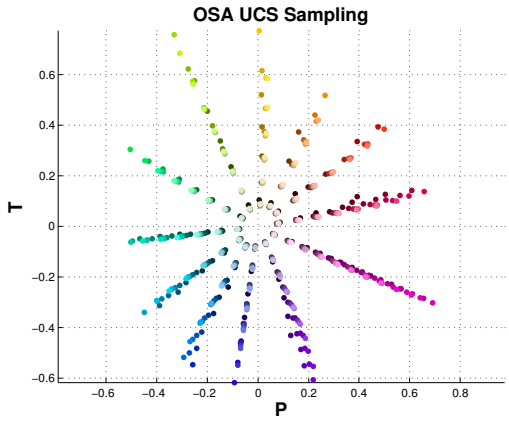
Hung-Berns varying lightness dataset

(b) CIECAM02 *JCh* space

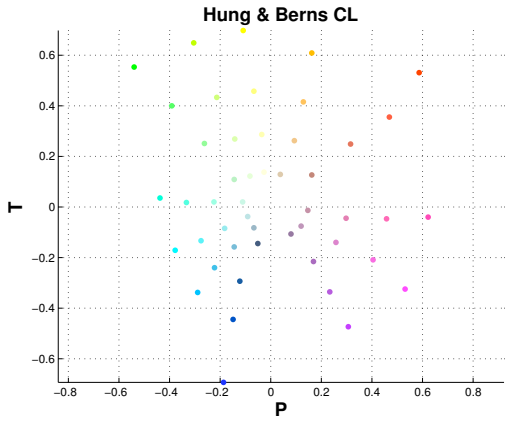
Figure 5.6. Constant hue datasets on different color spaces (continued, part 2 of 5).



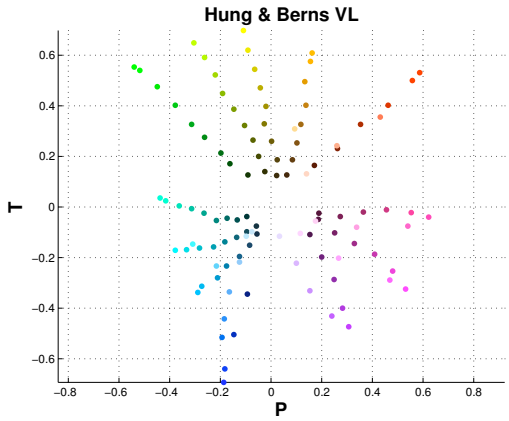
Ebner-Fairchild dataset



OSA UCS radial sampling



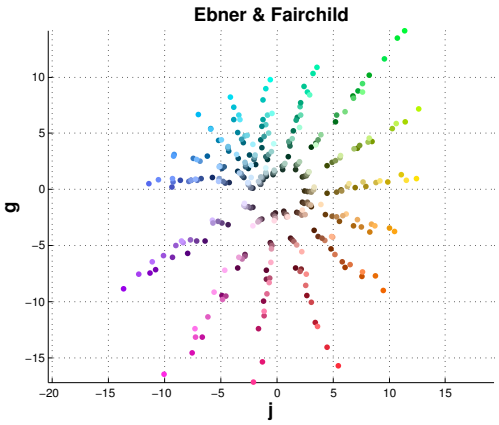
Hung-Berns constant lightness dataset



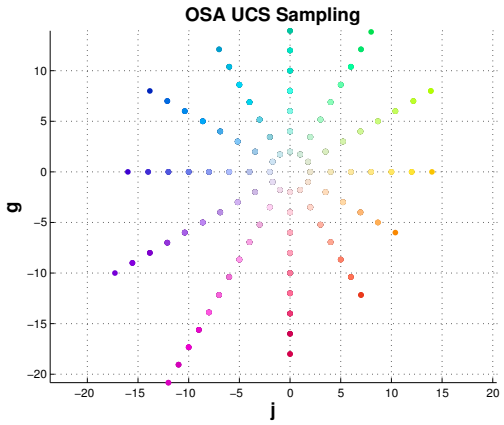
Hung-Berns varying lightness dataset

(c) IPT space

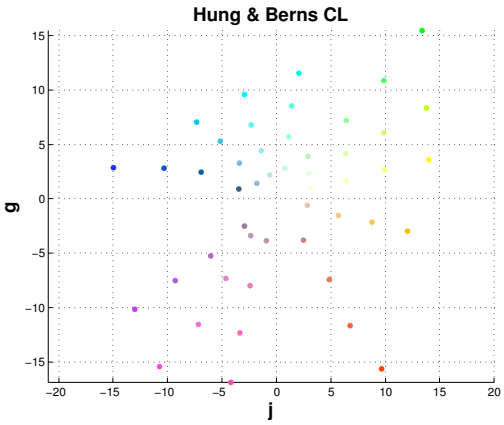
Figure 5.6. Constant hue datasets on different color spaces (continued, part 3 of 5).



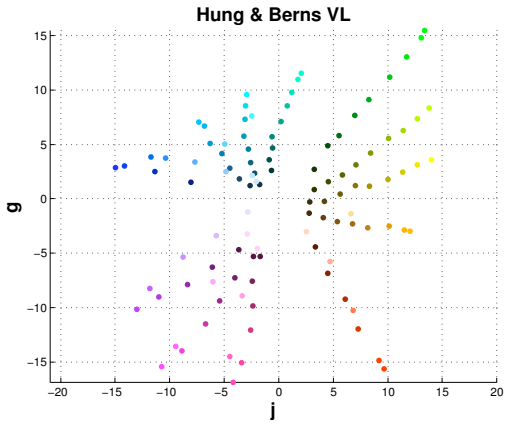
Ebner-Fairchild dataset



OSA UCS radial sampling



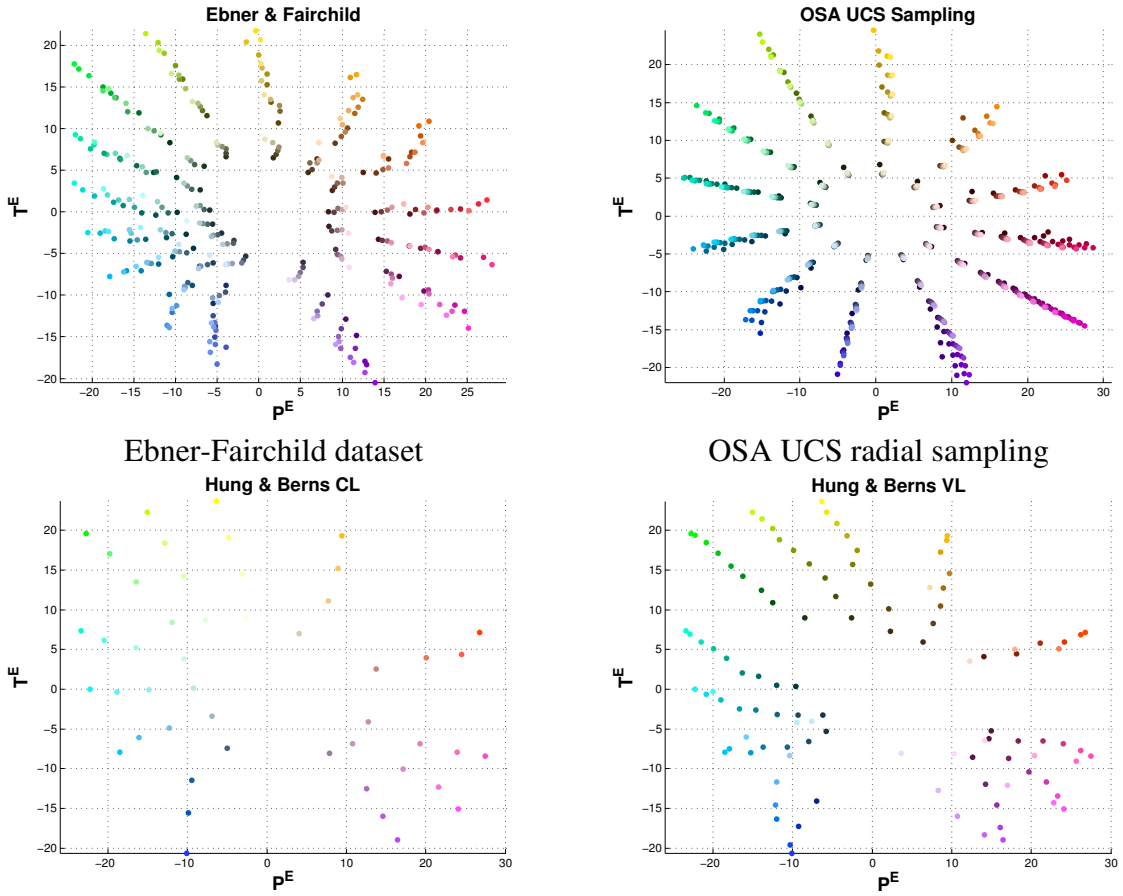
Hung-Berns constant lightness dataset



Hung-Berns varying lightness dataset

(d) OSA UCS *JLG* space

Figure 5.6. Constant hue datasets on different color spaces (continued, part 4 of 5).



Ebner-Fairchild dataset OSA UCS radial sampling

Hung-Berns constant lightness dataset Hung-Berns varying lightness dataset

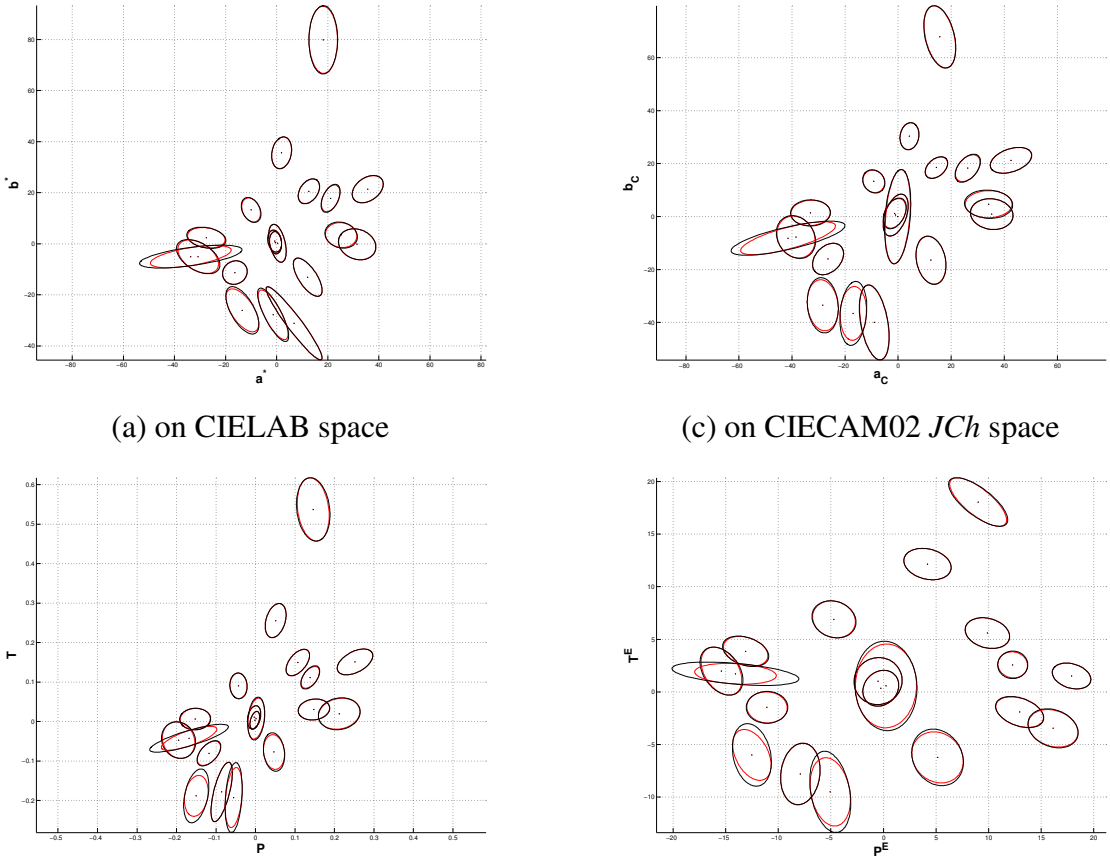
(e) IPT-EUC space

Figure 5.6. Constant hue datasets on different color spaces (continued, part 5 of 5).

5.7.4. Equal Color-Difference Ellipsoids

The equal color-difference ellipsoids of the RIT-DuPont dataset [Melgosa, 1997] were enlarged three times and plotted in the IPT-EUC, CIELAB, IPT and CIECAM02 *JCh* in Figure 5.7. For each color center, there are two ellipses representing the projection (black) and the intersection (red) of the equal color-difference ellipsoids on the chromatic plane. The tilt of the ellipsoids in the lightness direction can be represented by the mismatch of these two ellipses. In the IPT-EUC space, the neutral gray ellipsoid has circle projection and intersection with the chromatic plane. Comparing with these ellipses

in the CIELAB, IPT and CIECAM02 *JCh* spaces, the ellipses in the IPT-EUC space are more uniform and closer to round shape, which represented the better color uniformity. But, it was also found three ellipsoids of neutral color centers have different size, which indicates that there should be a lightness-dependent factor in chroma difference modeling in future color-difference formula development.



(a) on CIELAB space

(c) on CIECAM02 *JCh* space

(c) on IPT space

(d) on IPT-EUC space

Figure 5.7. Equal color-difference ellipsoids of the RIT-DuPont dataset on different color spaces.

5.7.5. Numerical Example

A numerical example of calculating the $I^EP^ET^E$, $\Delta E_{IPT\text{-}OPT}$, and $\Delta E^E_{IPT\text{-}EUC}$ is given in the Table 5.18. First, the tristimulus values were transformed to the values under 1931 2°

observer and CIE illuminant D65 using chromatic adaptation. Then, the IPT values were calculated based on the tristimulus values. The interim space I'P'T' were computed by applying the M_{COM}. Finally, $I^EP^ET^E$ were calculated by employing the lightness transformation and chromatic compression functions shown by Equations (5.13) and (5.14).

5.8. Conclusions

A Euclidean color space, IPT-EUC ($I^EP^ET^E$) was developed based on an optimized small color-difference formula $\Delta E_{IPT\text{-}OPT}$ on IPT space. IPT space provided a color appearance system with simple implementation and the theoretical basis of multi-stage color vision. Also, the non-uniform property of IPT space can be modeled by a simple form similar to CIE94 in chromatic plane, which is shown by the equal color-difference ellipsoids plotted in IPT space. A lightness transform function with three parameters was employed in the $\Delta E_{IPT\text{-}OPT}$ color-difference formula, which provided enough degrees of freedom to describe the non-uniformity of color difference along lightness. Furthermore, the IPT space was rotated to achieve two desired properties for uniform color spaces: The coordinates of the visually reddest patch in the Munsell color system were rotated onto the positive axis $+P^E$ of the IPT-EUC; in the middle of the IPT-EUC (medium gray), the visually equal color-differences along different directions on the chromatic plane have same numerical values, by which a standard was developed for comparing the color differences in other color regions. The performance evaluation was performed including STRESS value, F-test, hue constancy and equal color-difference ellipsoids. Furthermore, a deviation evaluation considering the visual uncertainty was performed. It was shown in the evaluations that the IPT-EUC space has quite good performance to evaluate color

quality and the Euclidean distance of the space $\Delta E^{E}_{IPT\text{-}EUC}$ was comparable to the best color-difference formula, CIEDE2000, available currently. Also, the IPT-EUC is capable to describe color as a uniform color space. The Euclidean color space IPT-EUC is a potential candidate of the unique color model for both color quality measuring and color describing.

6. Conclusion and Future Research

In this thesis, two topics in the color-difference formula and uniform color space modeling were recorded: to develop an evaluation metric on the performance of color-difference formula considering the visual uncertainty of the color-difference datasets and to derive a Euclidean color space based on IPT space for both color description and color quality measurement. Both of them were based on the step by step understanding on human color vision system and these datasets describing the system. Actually, most of the work in this thesis focused on analyzing the color-difference data. It is really a hard task to achieve color-difference data to represent human color vision properly since human's color vision is one the most complex system in the world: the system is multi-stages and non-linear; the system is really smart after thousand of year's evolution, which changes its status depending on different condition and environment; more important, there are quite a lot difference on color vision between individuals. Thus, when modeling these color-difference datasets, both the uncertainty of them and the difference between them cannot be neglected.

In the first topic, the uncertainty of the color-difference datasets was considered to evaluate the color-difference formulas. When there is uncertainty in these color-

difference datasets, it is not reasonable to model and evaluate the color-difference formula with the exact numbers. In this part of research, the color-difference datasets were transformed to a cloud instead of a number. The size of the cloud was determined by the uncertainty level in the visual experiments. Within the clouds, there are positions for every color-difference formula. Basically, if this position of one color-difference formula is not too close or too far away to the exact color-difference data (the most possible position of visual color difference), it is good to say that this color-difference formula fit the datasets properly. The evaluation metric was mathematically realized by using STRESS to describe the deviation. In addition, the fitting of the equal color-difference ellipsoids was included in the metric for these datasets proper to do so. Based on current understanding on color vision, the fitting of equal color-difference ellipsoids provided certain kinds of pre-process on the color-difference datasets to remove the errors caused by the visual experiment. But, it is still an open question that whether it is better to perform this pre-process since it will also remove certain characteristics of color vision shown by certain color-difference datasets. This evaluation metric provides another view on both the color-difference datasets developing and color-difference formulas modeling.

In the second topic, the difference between different datasets was the main consideration. Searching for a color model from these color-difference datasets was similar to combining different kinds of flowers to form a beautiful bouquet: remove the conflicts, find the harmonies and maximize different kinds of beauty. Even though they were performed following certain guidelines, these color-different datasets still have different experiment methodology, different experiment uncertainty and different

samples with different distributions in the color spaces. Thus, it is really difficult and impossible to model all of them with a simple color-difference formula or color space without considering all of these differences. Also, it is difficult to achieve precise results if the performance evaluation for color-difference formula and uniform color space was performed on a simple combination of these datasets without considering these differences. In this research, the efforts were applied on maximizing the usage of these differences during developing the color-difference formula and Euclidean color space. During the development of the color-difference formula and Euclidean color space, the four datasets were selectively used for deriving different kinds of coefficients based on the property of these datasets. A parametric factor k_L was also applied for describing different materials used in different datasets. In the performance evaluation, these four datasets were considered separately. In addition, to consider the differences between them, a step-wise optimization was used to extract as much as possible the similarity hidden in these datasets. Besides the datasets, another important aspect in this research is to find the proper form of the model. IPT space was selected because of its multi-stage color theory basis. The color-difference formula like CIE94 was selected since the Euclidean color space can be easily derived based on this form. The additional matrix and lightness functions were added as a bridge to connect the IPT space and the equation form like CIE94 after realizing the IPT space itself cannot be modeled well with this form. It was shown by the evaluation that the color model could be a potential candidate or basis of unique color model.

It seems that any future improvement of this research will be based on the deeper understanding of human color vision: both on how to achieving color-difference data and

how to construct the form the color-difference model. As usual, we do not fully understand these phenomena we are so familiar with.

Table 5.10. F-test results for the BFD-P dataset.

	CIELAB	CMC	CIE94	DIN99	CIEDE2000	DIN99d	CAM02-SCD	CAM02-UCS	OSA-UCS-GP	CAM02-OPT	CAM02-OPT-ROT	CAM02-EUC	CAM02-EUC_ROT	OSA-UCS-E	DIN99o	$\Delta E_{IPT\text{-}OPT}$	$\Delta E^{E}_{IPT\text{-}EUC}$
CIELAB	1.0000	**1.9432**	**1.4383**	**1.3413**	**2.0973**	**2.0231**	**1.9693**	**1.8165**	**1.9030**	**1.4245**	**1.2119**	**1.3483**	**1.1457**	**2.0791**	**2.0627**	**1.9973**	**2.0116**
CMC	***0.5146***	1.0000	***0.7402***	***0.6903***	1.0793	1.0411	1.0134	0.9348	0.9793	***0.7331***	***0.6237***	***0.6939***	***0.5896***	1.0699	1.0615	1.0279	1.0352
CIE94	***0.6953***	**1.3511**	1.0000	0.9326	**1.4582**	**1.4066**	**1.3692**	**1.2630**	**1.3231**	0.9904	***0.8426***	0.9375	***0.7966***	**1.4456**	**1.4341**	**1.3887**	**1.3986**
DIN99	***0.7456***	**1.4487**	1.0723	1.0000	**1.5636**	**1.5083**	**1.4682**	**1.3543**	**1.4188**	1.0620	***0.9035***	1.0053	***0.8542***	**1.5501**	**1.5378**	**1.4891**	**1.4997**
CIEDE2000	***0.4768***	0.9265	***0.6858***	***0.6395***	1.0000	0.9646	0.9390	***0.8661***	***0.9074***	***0.6792***	***0.5778***	***0.6429***	***0.5463***	0.9913	0.9835	0.9523	0.9591
DIN99d	***0.4943***	0.9605	***0.7109***	***0.6630***	1.0367	1.0000	0.9734	***0.8979***	0.9407	***0.7041***	***0.5990***	***0.6665***	***0.5663***	1.0277	1.0196	0.9873	0.9943
CAM02-SCD	***0.5078***	0.9867	***0.7304***	***0.6811***	1.0650	1.0273	1.0000	0.9224	0.9664	***0.7234***	***0.6154***	***0.6847***	***0.5818***	1.0558	1.0474	1.0142	1.0215
CAM02-UCS	***0.5505***	1.0698	***0.7918***	***0.7384***	**1.1546**	**1.1137**	1.0841	1.0000	1.0476	***0.7842***	***0.6672***	***0.7423***	***0.6307***	**1.1446**	**1.1355**	**1.0996**	**1.1074**
OSA-UCS-GP	***0.5255***	1.0211	***0.7558***	***0.7048***	**1.1021**	1.0631	1.0348	0.9545	1.0000	***0.7486***	***0.6368***	***0.7085***	***0.6021***	**1.0925**	1.0839	1.0496	1.0570
CIECAM02-OPT	***0.7020***	**1.3641**	1.0097	0.9416	**1.4723**	**1.4202**	**1.3824**	**1.2752**	**1.3359**	1.0000	***0.8507***	0.9465	***0.8043***	**1.4595**	**1.4480**	**1.4021**	**1.4121**
CIECAM02-OPT-ROT	***0.8252***	**1.6034**	**1.1868**	**1.1068**	**1.7306**	**1.6694**	**1.6250**	**1.4989**	**1.5703**	**1.1755**	1.0000	**1.1126**	0.9454	**1.7156**	**1.7020**	**1.6481**	**1.6599**
CIECAM02-EUC	***0.7417***	**1.4412**	1.0667	0.9948	**1.5555**	**1.5004**	**1.4605**	**1.3472**	**1.4114**	1.0565	***0.8988***	1.0000	***0.8497***	**1.5420**	**1.5298**	**1.4813**	**1.4919**
CIECAM02-EUC-ROT	***0.8728***	**1.6960**	**1.2553**	**1.1707**	**1.8305**	**1.7657**	**1.7188**	**1.5854**	**1.6610**	**1.2433**	1.0577	**1.1768**	1.0000	**1.8146**	**1.8003**	**1.7433**	**1.7557**
OSA-UCS-E	***0.4810***	0.9346	***0.6918***	***0.6451***	1.0087	0.9731	0.9472	***0.8737***	***0.9153***	***0.6852***	***0.5829***	***0.6485***	***0.5511***	1.0000	0.9921	0.9607	0.9675
DIN99o	***0.4848***	0.9421	***0.6973***	***0.6503***	1.0168	0.9808	0.9547	***0.8807***	0.9226	***0.6906***	***0.5875***	***0.6537***	***0.5555***	1.0080	1.0000	0.9683	0.9752
$\Delta E_{IPT\text{-}OPT}$	***0.5007***	0.9729	***0.7201***	***0.6715***	1.0500	1.0129	0.9860	***0.9095***	0.9528	***0.7132***	***0.6068***	***0.6751***	***0.5736***	1.0409	1.0327	1.0000	1.0071
$\Delta E^{E}_{IPT\text{-}EUC}$	***0.4971***	0.9660	***0.7150***	***0.6668***	1.0426	1.0057	0.9790	***0.9030***	0.9460	***0.7082***	***0.6025***	***0.6703***	***0.5696***	1.0336	1.0254	0.9929	1.0000

*$1/F_C = 1.0900$, $F_C = 0.9174$, 2072 color pairs.

Table 5.11. F-test results for the RIT-DuPont-Qiao dataset.

	CIELAB	CMC	CIE94	DIN99	CIEDE2000	DIN99d	CAM02-SCD	CAM02-UCS	OSA-UCS-GP	CAM02-OPT	CAM02-OPT-ROT	CAM02-EUC	CAM02-EUC_ROT	OSA-UCS-E	DIN99o	$\Delta E_{IPT\text{-}OPT}$	$\Delta E^{E}_{IPT\text{-}EUC}$
CIELAB	1.0000	**1.5722**	**1.9079**	**1.9512**	**2.3555**	**2.0938**	**1.7260**	**2.0782**	**1.9049**	**2.5479**	**2.8161**	**2.5668**	**2.7361**	**1.8185**	**1.7365**	**2.6999**	**2.6444**
CMC	***0.6360***	1.0000	1.2135	**1.2410**	**1.4982**	**1.3317**	1.0978	**1.3218**	1.2116	**1.6206**	**1.7911**	**1.6326**	**1.7402**	1.1566	1.1045	**1.7172**	**1.6820**
CIE94	***0.5241***	0.8241	1.0000	1.0227	**1.2346**	1.0974	0.9046	1.0892	0.9984	**1.3355**	**1.4760**	**1.3453**	**1.4341**	0.9531	0.9101	**1.4151**	**1.3860**
DIN99	***0.5125***	***0.8058***	0.9778	1.0000	1.2072	1.0731	0.8846	1.0651	0.9763	**1.3058**	**1.4433**	**1.3155**	**1.4022**	0.9320	0.8899	**1.3837**	**1.3553**
CIEDE2000	***0.4245***	***0.6675***	***0.8100***	0.8284	1.0000	0.8889	***0.7327***	0.8823	***0.8087***	1.0817	1.1956	1.0897	1.1616	***0.7720***	***0.7372***	1.1462	1.1227
DIN99d	***0.4776***	***0.7509***	0.9112	0.9319	1.1250	1.0000	0.8243	0.9925	0.9098	1.2169	**1.3450**	**1.2259**	**1.3067**	0.8685	0.8293	**1.2895**	**1.2630**
CAM02-SCD	***0.5794***	0.9109	1.1054	1.1305	**1.3647**	1.2131	1.0000	1.2041	1.1037	**1.4762**	**1.6316**	**1.4872**	**1.5852**	1.0536	1.0061	**1.5643**	**1.5322**
CAM02-UCS	***0.4812***	***0.7566***	0.9181	0.9389	1.1334	1.0075	0.8305	1.0000	0.9166	**1.2261**	**1.3551**	**1.2351**	**1.3166**	0.8750	0.8356	**1.2992**	**1.2725**
OSA-UCS-GP	***0.5250***	0.8254	1.0016	1.0243	**1.2365**	1.0992	0.9061	1.0910	1.0000	**1.3376**	**1.4784**	**1.3475**	**1.4363**	0.9546	0.9116	**1.4173**	**1.3882**
CIECAM02-OPT	***0.3925***	***0.6171***	***0.7488***	***0.7658***	0.9245	0.8218	***0.6774***	***0.8156***	***0.7476***	1.0000	1.1052	1.0074	1.0738	***0.7137***	***0.6815***	1.0596	1.0379
CIECAM02-OPT-ROT	***0.3551***	***0.5583***	***0.6775***	***0.6929***	0.8364	***0.7435***	***0.6129***	***0.7380***	***0.6764***	0.9048	1.0000	0.9115	0.9716	***0.6457***	***0.6166***	0.9587	0.9390
CIECAM02-EUC	***0.3896***	***0.6125***	***0.7433***	***0.7602***	0.9177	***0.8157***	***0.6724***	***0.8096***	***0.7421***	0.9926	1.0971	1.0000	1.0659	***0.7085***	***0.6765***	1.0518	1.0302
CIECAM02-EUC-ROT	***0.3655***	***0.5746***	***0.6973***	***0.7131***	0.8609	***0.7653***	***0.6308***	***0.7595***	***0.6962***	0.9312	1.0293	0.9381	1.0000	***0.6646***	***0.6347***	0.9868	0.9665
OSA-UCS-E	***0.5499***	0.8646	1.0492	1.0730	**1.2953**	1.1514	0.9491	1.1428	1.0475	**1.4011**	**1.5486**	**1.4115**	**1.5046**	1.0000	0.9549	**1.4847**	**1.4542**
DIN99o	***0.5759***	0.9054	1.0987	1.1237	**1.3565**	1.2058	0.9939	1.1968	1.0970	**1.4673**	**1.6217**	**1.4782**	**1.5756**	1.0472	1.0000	**1.5548**	**1.5229**
$\Delta E_{IPT\text{-}OPT}$	***0.3704***	***0.5823***	***0.7067***	***0.7227***	0.8724	***0.7755***	***0.6393***	***0.7697***	***0.7055***	0.9437	1.0430	0.9507	1.0134	***0.6735***	***0.6432***	1.0000	0.9795
$\Delta E^{E}_{IPT\text{-}EUC}$	***0.3782***	***0.5945***	***0.7215***	***0.7379***	0.8907	***0.7918***	***0.6527***	***0.7859***	***0.7203***	0.9635	1.0649	0.9706	1.0346	***0.6877***	***0.6566***	1.0210	1.0000

*$1/F_C = 1.2172$, $F_C = 0.8216$, 400 color vectors.

Table 5.12. F-test results for the Leeds dataset.

	CIELAB	CMC	CIE94	DIN99	CIEDE2000	DIN99d	CAM02-SCD	CAM02-UCS	OSA-UCS-GP	CAM02-OPT	CAM02-OPT-ROT	CAM02-EUC	CAM02-EUC_ROT	OSA-UCS-E	DIN99o	$\Delta E_{IPT\text{-}OPT}$	$\Delta E^{E}_{IPT\text{-}EUC}$
CIELAB	1.0000	**2.6444**	**1.7228**	**1.8169**	**4.3856**	**3.1219**	**3.3091**	**2.6726**	**2.1490**	**2.3692**	**1.8329**	**2.2084**	**1.7161**	**2.3511**	**2.8829**	**2.6923**	**2.7280**
CMC	***0.3782***	1.0000	***0.6515***	***0.6871***	**1.6584**	1.1806	1.2513	1.0106	0.8126	0.8959	***0.6931***	0.8351	***0.6489***	0.8891	1.0902	1.0181	1.0316
CIE94	***0.5804***	**1.5349**	1.0000	1.0546	**2.5456**	**1.8121**	**1.9207**	**1.5513**	1.2474	**1.3752**	1.0639	**1.2819**	0.9961	**1.3647**	**1.6734**	**1.5628**	**1.5835**
DIN99	***0.5504***	**1.4554**	0.9482	1.0000	**2.4138**	**1.7182**	**1.8213**	**1.4709**	1.1828	**1.3040**	1.0088	1.2155	0.9445	**1.2940**	**1.5867**	**1.4818**	**1.5015**
CIEDE2000	***0.2280***	***0.6030***	***0.3928***	***0.4143***	1.0000	***0.7118***	***0.7545***	***0.6094***	***0.4900***	***0.5402***	***0.4179***	***0.5036***	***0.3913***	***0.5361***	***0.6574***	***0.6139***	***0.6220***
DIN99d	***0.3203***	0.8471	***0.5519***	***0.5820***	**1.4048**	1.0000	1.0600	0.8561	***0.6884***	***0.7589***	***0.5871***	***0.7074***	***0.5497***	***0.7531***	0.9235	0.8624	0.8738
CAM02-SCD	***0.3022***	0.7991	***0.5206***	***0.5491***	**1.3253**	0.9434	1.0000	0.8076	***0.6494***	***0.7160***	***0.5539***	***0.6674***	***0.5186***	***0.7105***	0.8712	0.8136	0.8244
CAM02-UCS	***0.3742***	0.9895	***0.6446***	***0.6798***	**1.6410**	1.1681	1.2382	1.0000	0.8041	0.8865	***0.6858***	0.8263	***0.6421***	0.8797	1.0787	1.0074	1.0207
OSA-UCS-GP	***0.4653***	1.2306	0.8017	0.8455	**2.0408**	**1.4527**	**1.5399**	1.2437	1.0000	1.1025	0.8529	1.0277	0.7986	1.0941	**1.3415**	1.2529	**1.2695**
CIECAM02-OPT	***0.4221***	1.1162	***0.7272***	***0.7669***	**1.8511**	**1.3177**	**1.3967**	1.1280	0.9070	1.0000	***0.7736***	0.9321	***0.7243***	0.9924	1.2168	1.1364	1.1514
CIECAM02-OPT-ROT	***0.5456***	**1.4427**	0.9399	0.9913	**2.3927**	**1.7032**	**1.8054**	**1.4581**	1.1724	**1.2926**	1.0000	1.2049	0.9362	**1.2827**	**1.5729**	**1.4689**	**1.4883**
CIECAM02-EUC	***0.4528***	1.1974	***0.7801***	0.8227	**1.9859**	**1.4136**	**1.4984**	1.2102	0.9731	1.0728	0.8300	1.0000	***0.7771***	1.0646	**1.3054**	1.2191	1.2353
CIECAM02-EUC-ROT	***0.5827***	**1.5410**	1.0039	1.0588	**2.5556**	**1.8192**	**1.9283**	**1.5574**	1.2523	**1.3806**	1.0681	**1.2869**	1.0000	**1.3701**	**1.6800**	**1.5689**	**1.5897**
OSA-UCS-E	***0.4253***	1.1248	***0.7328***	***0.7728***	**1.8653**	**1.3278**	**1.4075**	1.1367	0.9140	1.0077	***0.7796***	0.9393	***0.7299***	1.0000	1.2262	1.1451	1.1603
DIN99o	***0.3469***	0.9173	***0.5976***	***0.6302***	**1.5212**	1.0829	1.1478	0.9270	***0.7454***	0.8218	***0.6358***	***0.7660***	***0.5952***	0.8155	1.0000	0.9339	0.9463
$\Delta E_{IPT\text{-}OPT}$	***0.3714***	0.9822	***0.6399***	***0.6748***	**1.6289**	1.1595	1.2291	0.9927	0.7982	0.8800	***0.6808***	0.8203	***0.6374***	0.8733	1.0708	1.0000	1.0133
$\Delta E^{E}_{IPT\text{-}EUC}$	***0.3666***	0.9694	***0.6315***	***0.6660***	**1.6076**	1.1444	1.2130	0.9797	***0.7877***	0.8685	***0.6719***	0.8095	***0.6290***	0.8618	1.0568	0.9869	1.0000

*$1/F_C = 1.2531$, $F_C = 0.7980$, 304 color pairs.

Table 5.13. F-test results for the Witt dataset.

	CIELAB	CMC	CIE94	DIN99	CIEDE2000	DIN99d	CAM02-SCD	CAM02-UCS	OSA-UCS-GP	CAM02-OPT	CAM02-OPT-ROT	CAM02-EUC	CAM02-EUC_ROT	OSA-UCS-E	DIN99o	$\Delta E_{IPT\text{-}OPT}$	$\Delta E^{E}_{IPT\text{-}EUC}$
CIELAB	1.0000	**1.9920**	**2.4803**	**2.1739**	**2.7441**	**2.8106**	**2.7065**	**2.7863**	**2.4176**	**2.5292**	**2.1077**	**2.5165**	**2.0728**	**2.0134**	**2.6505**	**2.4206**	**2.5997**
CMC	***0.5020***	1.0000	**1.2451**	1.0914	**1.3776**	**1.4110**	**1.3587**	**1.3988**	1.2137	**1.2697**	1.0581	**1.2633**	1.0406	1.0108	**1.3306**	1.2152	**1.3051**
CIE94	***0.4032***	***0.8031***	1.0000	0.8765	1.1064	1.1332	1.0912	1.1234	0.9747	1.0198	0.8498	1.0146	0.8357	***0.8118***	1.0686	0.9760	1.0481
DIN99	***0.4600***	0.9163	1.1409	1.0000	**1.2623**	**1.2928**	**1.2450**	**1.2817**	1.1121	1.1634	0.9696	1.1576	0.9535	0.9262	1.2192	1.1135	1.1958
CIEDE2000	***0.3644***	***0.7259***	0.9038	***0.7922***	1.0000	1.0242	0.9863	1.0154	0.8810	0.9217	***0.7681***	0.9170	***0.7554***	***0.7337***	0.9659	0.8821	0.9474
DIN99d	***0.3558***	***0.7087***	0.8825	***0.7735***	0.9764	1.0000	0.9630	0.9914	0.8602	0.8999	***0.7499***	0.8954	***0.7375***	***0.7164***	0.9431	0.8613	0.9250
CAM02-SCD	***0.3695***	***0.7360***	0.9164	***0.8032***	1.0139	1.0385	1.0000	1.0295	0.8933	0.9345	***0.7788***	0.9298	***0.7659***	***0.7439***	0.9793	0.8944	0.9605
CAM02-UCS	***0.3589***	***0.7149***	0.8902	***0.7802***	0.9849	1.0087	0.9713	1.0000	0.8677	0.9077	***0.7565***	0.9032	***0.7439***	***0.7226***	0.9513	0.8688	0.9330
OSA-UCS-GP	***0.4136***	0.8239	1.0259	0.8992	1.1351	1.1625	1.1195	1.1525	1.0000	1.0462	0.8718	1.0409	0.8574	0.8328	1.0963	1.0012	1.0753
CIECAM02-OPT	***0.3954***	***0.7876***	0.9806	0.8595	1.0850	1.1112	1.0701	1.1016	0.9559	1.0000	0.8333	0.9949	0.8195	***0.7961***	1.0479	0.9570	1.0278
CIECAM02-OPT-ROT	***0.4744***	0.9451	1.1767	1.0314	**1.3019**	**1.3334**	**1.2840**	**1.3219**	1.1470	1.2000	1.0000	1.1939	0.9834	0.9553	**1.2575**	1.1484	**1.2334**
CIECAM02-EUC	***0.3974***	***0.7916***	0.9856	0.8639	1.0905	1.1169	1.0755	1.1072	0.9607	1.0051	0.8376	1.0000	0.8237	***0.8001***	1.0533	0.9619	1.0331
CIECAM02-EUC-ROT	***0.4824***	0.9610	1.1966	1.0488	**1.3239**	**1.3559**	**1.3057**	**1.3442**	1.1663	1.2202	1.0169	1.2140	1.0000	0.9714	**1.2787**	1.1678	**1.2542**
OSA-UCS-E	***0.4967***	0.9893	**1.2319**	1.0797	**1.3629**	**1.3959**	**1.3442**	**1.3839**	1.2007	**1.2562**	1.0468	**1.2498**	1.0295	1.0000	**1.3164**	1.2022	**1.2912**
DIN99o	***0.3773***	***0.7515***	0.9358	0.8202	1.0353	1.0604	1.0211	1.0512	0.9121	0.9542	***0.7952***	0.9494	***0.7820***	***0.7596***	1.0000	0.9133	0.9808
$\Delta E_{IPT\text{-}OPT}$	***0.4131***	0.8229	1.0246	0.8981	1.1336	1.1611	1.1181	1.1511	0.9988	1.0449	0.8708	1.0396	0.8563	0.8318	1.0950	1.0000	1.0740
$\Delta E^{E}_{IPT\text{-}EUC}$	***0.3847***	***0.7662***	0.9541	0.8362	1.0556	1.0811	1.0411	1.0718	0.9300	0.9729	***0.8108***	0.9680	***0.7973***	***0.7745***	1.0196	0.9311	1.0000

*$1/F_C = 1.2295$, $F_C = 0.8133$, 362 color pairs.

Table 5.14. Performance evaluation with the visual uncertainty for the RIT-DuPont dataset.

Color pair groups	Percentage of F-Test Results for 5000 Randomized visual color difference comparing with Different Color-Difference Equations (Deviation to tolerance median visual color pairs)										Critical values at 95% Confidence Level		No. of Vectors
	CIELAB		CIE94		CIEDE2000		$\Delta E_{IPT\text{-}OPT}$		$\Delta E^{E}_{IPT\text{-}EUC}$				
	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*	
1	0.00	**100.00**	0.00	**99.94**	0.00	0.00	0.00	**100.00**	0.00	**100.00**	3.29	0.30	8
2	0.00	**100.00**	0.00	93.64	0.00	94.38	0.00	46.80	0.00	**99.72**	3.06	0.33	9
3	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	3.06	0.33	9
4	0.00	**100.00**	0.04	84.34	0.00	90.28	3.12	11.46	0.00	**100.00**	2.43	0.41	14
5	0.00	**100.00**	0.00	7.70	0.00	94.44	0.00	**100.00**	0.00	**99.96**	3.06	0.33	9
6	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**99.84**	1.02	0.00	3.06	0.33	9
7	0.00	**99.86**	0.00	0.00	0.00	0.00	0.00	**100.00**	0.00	0.00	3.29	0.30	8
8	0.00	**100.00**	0.00	**100.00**	0.00	0.00	0.00	**100.00**	0.00	**100.00**	3.06	0.33	9
9	0.00	98.12	0.04	38.50	0.00	70.50	1.02	12.42	0.38	18.70	3.06	0.33	9
10	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	3.60	0.28	7
11	0.00	**99.84**	0.00	88.40	0.00	90.18	0.00	96.82	0.00	**100.00**	4.04	0.25	6
12	0.00	**99.96**	0.00	**99.94**	0.00	**99.50**	0.00	**100.00**	0.00	**100.00**	3.60	0.28	7
13	0.00	**100.00**	0.00	95.86	0.00	99.12	0.10	35.14	0.00	93.08	4.04	0.25	6
14	0.00	56.58	0.00	50.84	0.00	37.48	0.00	63.00	0.00	74.08	3.60	0.28	7
15	0.00	64.96	1.62	49.00	0.34	56.02	0.00	94.22	0.00	99.30	3.60	0.28	7
16	0.00	93.22	0.00	57.78	0.96	31.18	0.00	**100.00**	0.00	**100.00**	3.60	0.28	7
17	0.00	98.50	0.00	98.94	***12.32***	4.16	0.00	**99.76**	0.00	**100.00**	3.60	0.28	7
18	0.00	**99.80**	0.00	**99.76**	0.00	**99.92**	0.00	**100.00**	0.00	**100.00**	3.60	0.28	7
19	0.00	97.88	0.00	67.18	***9.22***	13.90	0.08	37.28	3.40	21.70	3.60	0.28	7

Table 5.15. Performance evaluation with the visual uncertainty for the BFD-P dataset (1 of 4).

Color Area	Color pair groups	Percentage of F-Test Results for 5000 Randomized visual color difference comparing with Different Color-Difference Equations (Deviation to tolerance median visual color pairs)										Critical values at 95% Confidence Level		No. of Vectors	Average center of each color group				
		CIELAB		CIE94		CIEDE2000		$\Delta E_{IPT\text{-}OPT}$		$\Delta E^{E}_{IPT\text{-}EUC}$									
		Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*		*L**	*a**	*b**	*C**	*h*
Neutral	1	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	2.07	0.48	31	38.35	0.27	0.15	0.31	29.27
	2	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.62	0.62	68	50.34	-0.91	-0.58	1.08	-147.36
	3	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.35	0.74	174	62.75	0.28	0.36	0.46	52.27
	4	0.00	**100.00**	0.00	**100.00**	0.00	99.14	0.00	**100.00**	0.00	**100.00**	1.66	0.60	62	87.82	-1.41	2.35	2.74	120.95
Greenish	5	0.00	91.52	0.02	7.60	0.00	82.48	0.00	90.72	0.00	78.44	2.53	0.40	20	19.68	-17.46	1.68	17.54	174.52
	6	0.00	**99.76**	0.00	78.52	0.00	57.26	0.00	**99.98**	0.00	**99.92**	2.67	0.37	18	27.65	-29.04	-11.68	31.30	-158.09
	7	0.00	**99.98**	0.00	**99.80**	0.00	**99.98**	0.00	**100.00**	0.00	**99.98**	2.13	0.47	29	29.70	-17.42	-15.53	23.33	-138.28
	8	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	2.76	0.36	17	35.17	-43.74	3.76	43.90	175.09
	9	0.00	**100.00**	0.00	76.54	0.00	95.24	0.00	**99.84**	0.00	99.36	2.53	0.40	20	37.63	-30.92	-1.59	30.96	-177.06
	10	0.00	**100.00**	0.00	**100.00**	0.00	**99.82**	0.00	**100.00**	0.00	**99.94**	2.05	0.49	32	38.20	-26.79	22.53	35.00	139.94
	11	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.92	0.52	38	39.23	-22.69	-22.57	32.01	-135.16
	12	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	2.10	0.48	30	40.70	-44.56	36.63	57.68	140.58
	13	0.00	**99.90**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	2.10	0.48	30	42.48	-36.26	28.01	45.82	142.32
	14	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	2.13	0.47	29	47.91	-37.46	-21.52	43.21	-150.12
	15	0.00	**100.00**	0.10	71.80	0.18	66.10	0.10	71.78	0.26	61.18	2.10	0.48	30	47.94	-43.04	-1.09	43.05	-178.55
	16	0.00	98.18	0.00	61.00	0.00	61.32	0.00	52.00	0.00	39.92	2.86	0.35	16	54.69	-29.44	16.50	33.75	150.73
	17	0.00	98.34	0.28	23.94	0.34	22.30	0.00	88.76	0.00	50.74	2.46	0.41	21	55.31	-49.86	27.00	56.70	151.56
	18	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.35	0.74	169	56.34	-31.49	0.46	31.49	179.16
	19	0.00	44.98	0.00	35.42	0.00	34.04	0.00	45.52	0.00	30.48	7.15	0.14	6	56.85	-42.50	-7.00	43.07	-170.65
	20	0.00	56.94	0.00	30.28	0.00	13.04	0.06	4.28	0.16	2.70	7.15	0.14	6	60.88	-38.15	17.37	41.92	155.53
	21	0.00	**100.00**	0.00	**99.94**	0.00	**99.94**	0.00	**100.00**	0.00	**100.00**	1.86	0.54	42	60.94	-45.59	42.18	62.11	137.22
	22	0.00	91.58	0.00	17.24	0.00	25.08	0.00	92.28	0.00	82.28	4.03	0.25	10	62.72	-25.09	-7.85	26.29	-162.64
	23	0.00	57.92	0.04	17.14	0.22	5.92	0.00	37.72	0.02	18.96	3.47	0.29	12	63.09	-30.30	-4.75	30.67	-171.10

Table 5.15. Performance evaluation with the visual uncertainty for the BFD-P dataset (continued, 2 of 4).

Color Area	Color pair groups	Percentage of F-Test Results for 5000 Randomized visual color difference comparing with Different Color-Difference Equations (Deviation to tolerance median visual color pairs)										Critical values at 95% Confidence Level		No. of Vectors	Average center of each color group				
		CIELAB		CIE94		CIEDE2000		$\Delta E_{IPT\text{-}OPT}$		$\Delta E^{E}_{IPT\text{-}EUC}$									
		Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*		*L**	*a**	*b**	*C**	*h*
Greenish	24	0.00	17.14	0.00	97.16	0.00	90.00	0.00	29.82	0.00	18.62	2.98	0.34	15	63.17	-20.92	-15.13	25.82	-144.12
	25	0.00	50.04	0.00	95.46	0.00	69.28	0.02	31.04	0.04	27.46	2.36	0.42	23	64.13	-11.96	-7.53	14.13	-147.80
	26	0.00	54.50	0.00	81.00	0.00	58.08	0.00	59.38	0.00	48.78	3.28	0.31	13	64.25	-22.49	3.77	22.80	170.48
	27	0.00	**99.92**	0.00	**99.98**	0.00	98.46	0.00	98.10	0.00	95.36	1.75	0.57	51	71.55	-18.08	-0.66	18.09	-177.89
Bluish	28	0.00	37.40	0.00	8.30	0.00	31.32	0.00	95.28	0.00	94.28	4.03	0.25	10	14.00	-3.99	-5.53	6.83	-125.82
	29	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	98.64	0.00	98.40	2.23	0.45	26	22.46	17.76	-34.36	38.68	-62.67
	30	0.00	**100.00**	0.00	**100.00**	0.00	**99.68**	0.00	76.84	0.00	58.62	3.72	0.27	11	22.72	20.09	-46.69	50.83	-66.72
	31	0.00	38.08	0.00	87.94	0.66	13.74	0.00	55.22	0.00	54.66	2.86	0.35	16	23.04	-3.31	-6.29	7.11	-117.77
	32	0.00	96.94	0.00	98.56	0.00	**99.58**	0.00	98.78	0.00	98.16	2.23	0.45	26	24.57	-12.51	-14.25	18.96	-131.28
	33	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.92	0.52	38	29.69	8.36	-38.41	39.31	-77.72
	34	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	98.12	0.00	97.32	2.53	0.40	20	32.21	-4.23	-42.40	42.61	-95.70
	35	1.08	0.48	0.00	3.96	8.24	0.06	0.00	7.40	0.00	7.16	4.99	0.20	8	33.23	-14.70	-20.26	25.03	-125.97
	36	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.43	0.70	121	35.68	5.42	-31.04	31.51	-80.09
	37	0.00	34.50	0.00	92.26	0.00	62.84	0.00	**99.72**	0.00	96.80	2.98	0.34	15	37.44	-26.57	-29.04	39.36	-132.46
	38	0.00	**100.00**	0.00	**100.00**	0.00	82.74	0.00	48.76	0.00	45.66	1.75	0.57	51	41.86	2.56	-17.14	17.33	-81.51
	39	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	2.10	0.48	30	44.63	-30.44	-32.63	44.63	-133.01
	40	0.00	**100.00**	0.00	**100.00**	0.00	**99.92**	0.00	**99.58**	0.00	**99.58**	1.81	0.55	46	46.90	10.95	-14.67	18.31	-53.28
	41	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.76	0.57	50	47.25	-5.09	-24.66	25.18	-101.67
	42	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	2.10	0.48	30	51.37	5.60	-43.30	43.66	-82.63
	43	0.00	98.82	0.00	97.56	0.00	97.94	0.00	**99.92**	0.00	**99.92**	2.53	0.40	20	53.25	-0.69	-7.04	7.08	-95.56
	44	0.00	**99.92**	0.00	99.48	0.00	59.12	0.00	54.16	0.00	51.96	2.67	0.37	18	63.26	-3.06	-8.57	9.10	-109.63
	45	0.00	98.72	0.00	84.50	0.00	97.24	0.00	**100.00**	0.00	**100.00**	2.53	0.40	20	73.79	-14.10	-21.76	25.93	-122.94
	46	0.00	12.64	0.00	96.46	0.00	96.46	0.00	34.68	0.00	45.36	3.28	0.31	13	80.55	-13.38	-14.86	19.99	-132.00

Table 5.15. Performance evaluation with the visual uncertainty for the BFD-P dataset (continued, 3 of 4).

Color Area	Color pair groups	Percentage of F-Test Results for 5000 Randomized visual color difference comparing with Different Color-Difference Equations (Deviation to tolerance median visual color pairs)										Critical values at 95% Confidence Level		No. of Vectors	Average center of each color group				
		CIELAB		CIE94		CIEDE2000		$\Delta E_{IPT\text{-}OPT}$		$\Delta E^{E}_{IPT\text{-}EUC}$									
		Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*		*L**	*a**	*b**	*C**	*h*
Bluish	47	0.00	66.78	0.00	69.84	0.00	52.16	0.00	79.78	0.00	75.04	3.72	0.27	11	83.11	-4.16	-7.95	8.97	-117.65
Reddish	48	0.00	97.84	0.00	**100.00**	0.00	98.80	0.00	98.26	0.00	97.98	2.53	0.40	20	16.39	20.79	-11.99	24.00	-29.96
	49	0.00	94.66	0.00	81.48	0.00	37.20	0.00	91.98	0.00	86.14	3.47	0.29	12	22.38	12.47	7.38	14.50	30.63
	50	0.00	**100.00**	0.00	99.32	0.00	**99.74**	0.00	**100.00**	0.00	**100.00**	2.10	0.48	30	32.92	59.29	-0.07	59.30	-0.06
	51	0.00	**100.00**	0.00	90.80	0.00	**99.90**	0.00	97.64	0.00	98.90	2.53	0.40	20	34.15	54.16	43.68	69.58	38.89
	52	0.00	98.40	0.04	15.02	0.00	28.66	0.00	85.46	0.00	73.82	4.43	0.23	9	34.62	31.20	-10.42	32.89	-18.47
	53	0.00	**100.00**	0.00	**99.94**	0.00	**99.66**	0.00	98.78	0.00	98.66	2.13	0.47	29	34.87	27.25	-21.10	34.46	-37.76
	54	0.00	**100.00**	0.00	90.20	0.00	99.02	0.00	**100.00**	0.00	**99.96**	2.60	0.39	19	36.92	47.13	17.98	50.44	20.88
	55	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	84.56	0.00	91.60	2.41	0.42	22	39.33	52.89	48.03	71.44	42.24
	56	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	100.00	1.38	0.73	153	44.76	37.82	23.22	44.39	31.55
	57	0.00	**100.00**	0.00	99.34	0.00	**99.86**	0.00	**99.84**	0.00	**99.74**	2.10	0.48	30	45.44	64.63	46.82	79.80	35.92
	58	0.00	9.92	0.00	15.28	0.00	2.82	0.00	51.80	0.00	24.56	5.82	0.17	7	50.13	25.39	2.75	25.54	6.17
	59	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.85	0.54	43	54.79	51.47	36.49	63.09	35.34
	60	0.00	59.22	0.00	**100.00**	0.00	**100.00**	0.00	73.78	0.00	65.44	2.00	0.50	34	60.93	26.05	-14.39	29.77	-28.92
	61	0.00	97.04	0.00	6.22	0.00	9.86	0.00	93.62	0.00	83.94	3.47	0.29	12	62.38	23.17	18.42	29.60	38.48
	62	0.00	90.52	0.00	99.00	0.00	88.00	0.04	55.88	0.64	30.80	2.19	0.46	27	64.07	33.70	2.63	33.80	4.47
	63	0.00	77.04	0.00	63.34	0.20	3.28	0.00	56.46	0.00	35.28	2.36	0.42	23	65.93	14.10	4.62	14.84	18.16
Yellowish	64	0.00	99.38	0.00	**99.70**	0.00	**99.98**	0.00	**100.00**	0.00	**100.00**	2.13	0.47	29	31.93	7.79	30.17	31.16	75.53
	65	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.62	0.62	69	52.42	42.02	53.08	67.70	51.64
	66	0.00	**100.00**	0.00	97.48	0.00	74.84	0.00	65.22	0.00	47.92	2.76	0.36	17	59.41	-18.58	57.71	60.63	107.85
	67	0.00	78.66	0.10	1.78	0.10	1.74	0.00	16.94	0.04	4.64	3.47	0.29	12	60.14	-25.30	38.22	45.83	123.50
	68	0.00	87.58	0.00	36.78	0.00	38.60	0.00	49.52	0.00	22.42	4.03	0.25	10	60.26	-34.01	36.27	49.72	133.16

Table 5.15. Performance evaluation with the visual uncertainty for the BFD-P dataset (continued, 4 of 4).

Color Area	Color pair groups	Percentage of F-Test Results for 5000 Randomized visual color difference comparing with Different Color-Difference Equations (Deviation to tolerance median visual color pairs)										Critical values at 95% Confidence Level		No. of Vectors	Average center of each color group				
		CIELAB		CIE94		CIEDE2000		$\Delta E_{IPT\text{-}OPT}$		$\Delta E^{E}_{IPT\text{-}EUC}$									
		Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*		*L**	*a**	*b**	*C**	*h*
Yellowish	69	0.00	77.90	0.00	23.36	0.12	9.76	0.00	17.16	0.00	17.58	4.03	0.25	10	60.75	-7.83	17.19	18.89	114.48
	70	0.00	**100.00**	0.00	**99.98**	0.00	97.60	0.00	92.92	0.00	82.58	3.28	0.31	13	61.14	1.31	65.77	65.78	88.86
	71	0.00	**99.98**	0.00	**99.58**	0.00	75.58	0.12	8.66	0.02	31.32	2.98	0.34	15	61.61	20.47	62.94	66.18	71.98
	72	0.00	**100.00**	0.00	95.94	0.00	79.00	0.00	48.82	0.00	19.94	2.67	0.37	18	61.87	-10.46	40.94	42.25	104.33
	73	0.00	**99.70**	0.00	**100.00**	0.00	98.56	0.00	86.24	0.00	84.60	2.76	0.36	17	62.03	10.77	29.84	31.73	70.16
	74	0.16	19.34	0.00	77.84	0.02	35.02	0.00	90.42	0.00	84.84	2.98	0.34	15	62.48	8.57	9.13	12.52	46.81
	75	0.00	47.56	0.02	1.22	0.00	6.80	0.00	10.20	0.00	15.70	4.43	0.23	9	62.60	1.88	20.95	21.04	84.87
	76	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.70	0.59	57	64.97	30.78	35.47	46.96	49.06
	77	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	2.10	0.48	30	70.05	-2.03	43.39	43.44	92.68
	78	0.00	99.08	0.00	92.84	0.00	92.40	0.00	70.14	0.00	75.38	3.12	0.32	14	71.28	-7.05	44.79	45.34	98.95
	79	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.70	0.59	57	75.91	-3.50	84.52	84.59	92.37
	80	0.00	**100.00**	0.00	**100.00**	0.00	**99.34**	0.00	97.80	0.00	**99.52**	2.53	0.40	20	83.55	2.93	102.10	102.15	88.36
	81	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.45	0.69	112	86.86	-6.23	46.86	47.27	97.57

Table 5.16. Performance evaluation with the visual uncertainty for the Leeds dataset.

Color Area	Color pair groups	Percentage of F-Test Results for 5000 Randomized visual color difference comparing with Different Color-Difference Equations (Deviation to tolerance median visual color pairs)										Critical values at 95% Confidence Level		No. of Vectors	Average center of each color group				
		CIELAB		CIE94		CIEDE2000		$\Delta E_{IPT\text{-}OPT}$		$\Delta E^{E}_{IPT\text{-}EUC}$									
		Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*		*L**	*a**	*b**	*C**	*h*
Neutral	1	0.00	11.42	0.00	11.42	0.00	9.42	0.00	9.90	0.00	9.90	3.72	0.27	11	28.12	-0.16	0.57	0.60	105.79
	2	2.98	0.16	2.98	0.16	2.20	0.24	2.68	0.16	2.68	0.16	4.43	0.23	9	39.81	-0.21	0.34	0.40	121.27
	3	0.00	**100.00**	0.00	**100.00**	0.00	**99.90**	0.00	**100.00**	0.00	**100.00**	1.72	0.58	54	49.82	-0.35	0.53	0.63	123.79
	4	0.00	67.22	0.00	67.22	0.00	44.44	0.00	36.92	0.00	36.92	2.60	0.39	19	63.57	-0.25	-0.03	0.25	-173.17
	5	0.00	25.76	0.00	25.76	0.00	23.94	0.00	23.56	0.00	23.56	4.03	0.25	10	78.21	-0.35	-0.50	0.61	-125.01
	6	0.00	45.80	0.00	45.84	0.00	38.44	0.00	39.82	0.00	39.94	4.43	0.23	9	87.96	-0.48	-0.44	0.65	-137.76
Greenish	7	0.00	**100.00**	0.00	54.58	0.00	57.86	0.00	75.18	0.00	76.78	3.12	0.32	14	55.00	-31.85	0.06	31.85	179.98
Bluish	8	0.00	**100.00**	0.00	**100.00**	0.00	97.70	0.00	**99.88**	0.00	**99.96**	1.65	0.61	64	36.32	-4.30	-28.02	28.35	-98.77
Reddish	9	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.74	0.57	52	44.38	36.72	22.01	42.81	30.96
Yellowish	10	0.00	**100.00**	0.00	93.56	0.00	40.24	0.00	40.94	0.00	33.58	2.98	0.34	15	85.91	-6.41	45.82	46.26	98.02

Table 5.17. Performance evaluation with the visual uncertainty for the Witt dataset.

Color Area	Color pair groups	Percentage of F-Test Results for 5000 Randomized visual color difference comparing with Different Color-Difference Equations (Deviation to tolerance median visual color pairs)										Critical values at 95% Confidence Level		No. of Vectors	Average center of each color group				
		CIELAB		CIE94		CIEDE2000		$\Delta E_{IPT\text{-}OPT}$		$\Delta E^{E}_{IPT\text{-}EUC}$									
		Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	1/*Fc*	*Fc*		*L**	*a**	*b**	*C**	*h*
Neutral	1	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.54	0.65	85	62.70	0.11	0.24	0.27	66.55
Greenish	2	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.57	0.64	78	56.22	-31.21	0.48	31.21	179.21
Bluish	3	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.54	0.65	85	35.57	5.25	-31.35	31.79	-80.54
Reddish	4	0.00	**100.00**	0.00	89.90	0.00	99.46	0.00	98.44	0.00	96.30	1.54	0.65	85	44.59	37.18	23.15	43.80	31.92
Yellowish	5	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	0.00	**100.00**	1.55	0.65	83	86.84	-6.96	46.60	47.12	98.55

Table 5.18. Numerical examples.

	C, 1931 2° observer						D65, 1931 2° observer						*I*	*P*	*T*	I^E	P^E	T^E	$\Delta E_{ab}{}^{*}$	$\Delta E_{IPT\text{-}OPT}$	$\Delta E^{E}_{IPT\text{-}EUC}$
	x	*y*	*Y*	*X*	*Y*	*Z*	*X*	*Y*	*Z*	*L**	*a**	*b**									
Sample1	0.5734	0.3057	12.00	22.51	12.00	4.75	22.25	11.95	4.36	41.13	61.87	30.09	0.3784	0.4618	0.2111	44.12	24.56	0.00	3.88	0.654	0.722
Sample2	0.5821	0.3125	12.30	22.91	12.30	4.15	22.65	12.25	3.81	41.61	61.67	33.93	0.3791	0.4599	0.2328	44.21	24.50	0.71			

White Point	X	Y	Z
C, 1931 2 observer	98.07	100.000	118.23
D65, 1931 2 observer	95.047	100.000	108.883

7. References

[Alder, 1981]

Alder C. A Monte Carlo method for the validation of discrimination ellipse data. J Soc Dyers Colour 1981;97:514-517.

[Alder, 1982]

Alder C, Chaing KP, Chong TF, Coates E, Khalili AA, Rigg B, Uniform Chromaticity scales – new experimental data, J Soc Dyers Colour, 1982;98:14-20.

[Berns, 1991]

Berns RS, Alman DH, Reniff L, Snyder GD, Balonon-Rosen MR. Visual determination of suprathreshold color-difference tolerances using probit analysis. Color Res Appl 1991;16:297-316.

[Berns, 1993]

Berns RS. The mathematical development of CIE TC 1-29 proposed color-difference equation: CIELCH. Proceedings of AIC Colour 93, Budapest, Hungary: ISCC 1993. Volume B, p189-192.

[Berns, 2000]

Berns RS, Billmeyer and Saltzman's Principles of Color Technology, 3rd edition, New York, NY, USA: John Wiley; 2000.

[Berns, 2002]

Berns RS, Derivation of a hue-angle dependent, hue-difference weighting function for CIEDE2000, SPIE Proceedings 4421, 9th Congress of the International Colour Association, 638-641, 2002.

[Berns, 2007]

Berns RS, Xue Y. Optimizing color-difference equations and uniform color spaces for industrial tolerancing. Proceedings of AIC 2007, Hangzhou, China: ISCC 2007. p 24-28.

[Berns, 2008]

Berns RS, Generalized industrial color-difference space based on multi-stage color-vision and line-element integration, Óptica Pura Y Aplicada 41 pp. 301-311, 2008.

[Balonon-Rosen, 1993]

Balonon-Rosen MR, An Uncertainty Analysis of A Color Tolerance Database, M.S. Thesis Rochester, NY: Rochester Institute of Technology, 1993.

[Cheung, 1986]

Cheung M and Rigg B, Colour-difference ellipsoids for five CIE colour centres, Color Res Appl 1986;11:185-195.

[CIE Publ. 131, 1998]

CIE, The CIE1997 Interim Colour Appearance Model (Simple Version), CIECAM97s, CIE Pub. 131, 1998.

[CIE Publ. 142-2001, 2001]

CIE, Improvement to Industrial Colour-Difference Evaluation, CIE Publ. 142-2001, 2001

[CIEPubl. 15, 2004]

CIE, Colorimetry, third edition, CIE Pub. 15, 2004.

[CIE Publ. 159, 2004]

CIE, A Colour Appearance Model for Colour Management Applications: CIECAM02, CIE Pub. 159, 2004.

[Clarke, 1984]

Clarke FJJ, McDonald R, Rigg B. Modification to the JPC79 colour-difference formula. J Soc Dyers Colour 1984;100:128-132.

[Coates, 1981]

Coates E, Fong KY, and Rigg B, Uniform lightness scales. J Soc Dyers Colour 1981; 97:179 – 183.

[Cui, 2002]

Cui G, Luo MR, Rigg B, Roesler G, Witt K. Uniform colour spaces based on the DIN99 colour-difference formula. Color Res Appl 2002;27:282-290.

[DIN6176, 2000]

DIN 6176: Farbmetrische Bestimmung von Farbabständen bei Köroerfarben nach der DIN99-Formel. Berlin: DIN Deutsche Institut für Normung e.V. 2000.

[Ebner, 1998a]

Ebner F, Fairchild MD. Finding constant hue surfaces in color space, Proceedings of SPIE 3300, 1998;3300:107-117.

[Ebner, 1998b]

Ebner F, Fairchild MD. Development and testing of a color space (IPT) with improved hue uniformity. Proceedings of the 6th Color Imaging Conference, Scottsdale, AZ: IS&T; 1998 p 8-13.

[Fairchild, 1993]

Fairchild MD, Berns RS. Image color appearance specification through extension of CIELAB. Color Res Appl 1993;18:178-190.

[Fairchild, 1996]

Fairchild MD. Refinement of the RLAB color space. Color Res Appl 1996;21:338-346.

[Fairchild, 2005]

Fairchild MD. Color Appearance Models, 2nd edition. Chichester, NJ: John Wiley; 2005.

[Garcia, 2007]

Garcia PA, Huertas R, Melgosa M, Cui G. Measurement of the relationship between perceived and computed color differences. J Opt Soc Am A 2007;24: 1823-1829.

[Green, 2002]

Green P and MacDonald L, Color Engineering: Achieving Device Independent Colour, Chichester: John Wiley; 2002.

[Guan, 1999]

Guan SS, Luo MR. Investigation of parametric effects using small colour differences. Color Res Appl 1999;24:331-343.

[Guild, 1931]

Guild J, The colorimetric properties of the spectrum, Philos. Trans. Roy. Soc. London, 230 (A), pp. 149-187, 1931.

[Hård, 1996a]

Hård A, Sivik L, and Tonnquist G, NCS natural color system-from concept to research and application, Part I, Color Res Appl 1996;21:180-205.

[Hård, 1996b]

Hård A, Sivik L, and Tonnquist G, NCS natural color system-from concept to research and application, Part II, Color Res Appl 1996;21:206-220.

[Huertas, 2006]

Huertas R, Melgosa M, Oleari C. Performance of a color-difference formula based on OSA-UCS space using small-medium color differences. J Opt Soc Am A 2006;23:2077-2084.

[Hung, 1995]

Hung PC, Berns RS. Determination of constant hue loci for a CRT gamut and their predictions using color appearance spaces. Color Res Appl 1995;20:285-295.

[Hunt, 1991]

Hunt RWG. Revised colour-appearance model for related and unrelated colours. Color Res Appl 1991;16:146-165.

[Hunt, 1994]

Hunt RWG. An improved predictor of colourfulness in a model of colour vision. Color Res Appl 1994;19:23-26.

[Hunter, 1987]

Hunter RS and Harold RW, The measurement of appearance, New York, NY, USA: John Wiley; 1987.

[ICC, 1996]

ICC, A Standard Default Color Space for the Internet: sRGB, Version 1.10, ICC, 1996.

[ICC specification, 2004]

ICC, Specification ICC. 1: 2004-10 (Profile version 4.2.0.0), ICC, 2004.

[Indow, 1991]

Indow T and Morrison ML, Construction of discrimination ellipsoids for surface colors by the method of constant stimuli, Color Res Appl 1991;16:42-56.

[Jain, 1988]

Jain AK, Dubes RC. Algorithms for Clustering Data. Englewood Cliffs, NJ: Prentice-Hall; 1988.

[Judd, 1936]

Judd DB, Estimation of chromaticity differences and nearest color temperature on the standard 1931 ICI colorimetric coordinate system, J Opt Soc Am 1936;26:421-426.

[Kim, 1997a]

Kim DH, Nobbs JH. New weighting functions for the weighted CIELAB colour difference formulae. Proceedings of AIC Colour 97, Kyoto, Japan: ISCC 1997. p 446-449.

[Kim, 1997b]

Kim DH, The Influence of Parametric Effects on the Appearance of Small Colour Differences, Ph. D. Thesis, Leeds, UK: University of Leeds; 1997.

[Kim, 2001]

Kim DH, Cho EK, Kim JP. Evaluation of CIELAB-based colour-difference formulae using a new dataset. Color Res Appl 2001;26:369-375.

[Kuehni, 1999]

Kuehni RG, Towards an improved uniform color spaces, Color Res Appl 1999;24:253-265.

[Kuehni, 2008]

Kuehni RG, Color difference formulas: An unsatisfactory state of affairs, Color Res Appl 2008;33:324-326.

[Kurzhanskiy, 2006]

Kurzhanskiy AA, http://www.eecs.berkeley.edu/~akurzhan/ellipsoids/. Ellipsoidal Toolbox, Matlab Toolbox.

[Li, 2002]

Li C, Luo MR. A uniform color space based upon CIECAM97s. Proceedings of SPIE 2002;4421:579-582.

[Li, 2003]

Li C, Luo MR, Cui G. Colour-differences evaluation using colour appearance models. Proceedings of the 11th Color Imaging Conference. Scottsdale, AZ: IS&T; 2003. p 127-131.

[Luo, 1986]

Luo MR, Rigg B. Chromaticity-discrimination ellipses for surface colours. Color Res Appl 1986;11: 25-42.

[Luo, 1987a]

Luo MR and Rigg B, BFD (l:c) colour-difference formula Part I – development of the formula. J Soc Dyers Colour 1987;103:86-94.

[Luo, 1987b]

Luo MR, and Rigg B, BFD(l:c) colour-difference formula, Part II-Performance of the formula. J Soc Dyers Colour 1987;103:126 – 132.

[Luo, 1996]

Luo MR, Lo MC and Kuo WG, The LLAB(l:c) colour model, Color Res Appl 1996;21:412-428.

[Luo, 1999]

Luo MR, Personal communication, 1999.

[Luo, 2001]

Luo MR, Cui G, and Rigg B, The development of the CIE 2000 colour-difference formula: CIEDE2000. Color Res Appl 2001;26:340-350.

[Luo, 2006]

Luo MR, Cui G, Li C. Uniform colour spaces based on CIECAM02 colour appearance model. Color Res Appl 2006;31:320-330.

[MacAdam, 1942]

MacAdam, DL, Visual sensitivities to color difference in daylight. J Opt Soc Am 1942;32:247-274.

[MacAdam, 1974]

MacAdam DL, Uniform color scales, J Opt Soc Am 1974;64:1691-1702.

[Maier, 1992]

Maier T, Written communication to CIE TC 1-29, 1992.

[Melgosa, 1997]

Melgosa M, Hita E, Poza AJ, Alman DH, Berns RS. Suprathreshold color-difference ellipsoids for surface colors. Color Res Appl 1997;22:148-155.

[Melgosa, 2004]

Melgosa M, Huertas R. Relative significance of the terms in the CIEDE2000 and CIE94 color-difference formulas. J Opt Soc Am A 2004;21:2269-2275.

[Melgosa, 2008]

Melgosa M, Huertas R, Berns RS. Performance of recent advanced color-difference formulae using the Standardized Residual Sum of Squares index. J Opt Soc Am A 2008;25:1828-1834.

[Moroney, 2003]

Moroney N, A radial sampling of the OSA uniform color scales, Proceedings of the 11th Color Imaging Conference. Scottsdale, AZ: IS&T; 2003, p 175-180.

[Nayatani, 1990]

Nayatani Y, Takahama K, Sobagaki H, Hashimoto K. Color-appearance model and chromatic adaptation transform. Color Res Appl 1990;15:210-221.

[Oleari, 2004]

Oleari C. Color opponencies in the system of the uniform color scales of the Optical Society of America. J Opt Soc Am A 2004;21:677-682.

[Oleari, 2008]

Oleari C, Melgosa M and Huertas R, Euclidean color-difference formula for small-medium color differences in log-compressed OSA-UCS space, manuscript, 2008.

[Pointer, 1980]

Pointer MR. The gamut of real surface colours. Color Res Appl 1980;5:145-155.

[Qiao, 1998]

Qiao Y, Berns RS, Reniff L, and Montag E. Visual determination of hue suprathreshold color-difference tolerance. Color Res Appl 1998;23**:**302-313.

[Reniff, 1989]

Reniff LA, Visual Determination of Color Differences Using Probit AnalysisL Phase II, M. S. Thesis, Rochester, NY: Rochester Institute of Technology, 1989.

[Rich, 1983]

Rich DC and Billmeyer Jr. FW, Small and moderate color differences, IV: Color-difference-perceptibility ellipses in surface-color space, Color Res Appl 1983;8:31-39.

[Robertson, 1978]

Robertson AR. CIE guidelines for coordinated research on colour-difference evaluation. Color Res Appl 1978;3:149-151.

[Schultz, 1972]

Schultz W, The usefulness of colour-differences formulae for fixing colour tolerances, colour metrics. Soesterberg: AIC, Holland: pp. 245 – 265, 1972.

[Snyder, 1991]

Snyder GD, Visual Determination of Industrial Color-Difference Tolerances Using Probit Analysis, M. S. Thesis, Rochester, NY: Rochester Institute of Technology, 1991.

[Shamey, 2008]

Shamey R, Lee SG, Hinks D, Jasper W, Performance of recent color difference formulas around a CIE blue color center, 4th European Conference on Colour in Graphics, Imaging, and Vision, 10th International Symposium on Multispectral Colour Science. Terrassa, Spain: IS&T; pp 7-11, 2008.

[Strocka, 1983]

Strocka D, Brockes A, Paffhausen W. Influence of experimental parameters on the evaluation of color-difference ellipsoids. Color Res Appl 1983;8:169-175.

[Stromeyer III, 1985]

Stromeyer III CF, Cole GR, Kronauer RE. Second-site adaptation in the red-green chromatic pathways. Vision Res 1985;25:219-237.

[Thomsen, 2000]

Thomsen K. A Euclidean color space in high agreement with the CIE94 color difference formula. Color Res Appl 2000;25:64-65.

[Urban, 2007a]

Urban P, Berns RS, Rosen MR. Constructing Euclidean color spaces based on color difference formulas. Proceedings of the 15th Color Imaging Conference, Albuquerque, NM: IS&T; 2007 p 77-81.

[Urban, 2007b]

Urban P, Rosen MR, Berns RS, Schleicher D. Embedding non-euclidean color spaces into euclidean color spaces with minimal isometric disagreement. J Opt Soc Am A 2007;24:1516-1528.

[von Kries, 1902]

von Kries J, Chromatic adaptation, Fribourg, German: Festschrift der Albercht-Ludwig-Universität, 1902 (Translation: MacAdam DL, Sources of Color Science, Cambridge, MA: MIT Press, 1970).

[von Seggern, 1990]

von Seggern DH. CRC Handbook of Mathematical Curves and Surfaces. Boca Raton, Florida: CRC Press; 1990.

[Völz, 1998]

Völz HG, Transformation der CIE94-Formel in einen euklidischen Farbraum. Die Farbe 1998;44:97-105.

[Völz, 1999-2000]

Völz HG. Die Euklidisierung des CMC-Raumes zur Berechnung grosser Farbabstände. Die Farbe 1999-2000;45:1-23.

[Völz, 2006]

Völz HG. Euclidization of the first quadrant of the CIEDE2000 color difference system for the calculation of large color differences. Color Res Appl 2006;31:5-12.

[Witt, 1987]

Witt K. Three-dimensional threshold of color-difference perceptibility in painted samples: variability of observers in four CIE color regions. Color Res Appl 1987;12:128-134.

[Witt, 1990]

Witt K, Parametric effects on surface color-difference evaluation at threshold. Color Res Appl 1990;15:189-199.

[Witt, 1995]

CIE guidelines for coordinated future work on industrial colour-difference evaluation, Color Res Appl 1995;20:399-403.

[Witt, 1999]

Witt K. Geometric relations between scales of small colour differences. Color Res Appl 1999;24:78-92.

[Witt, 2009]

Personal Communication, 2009

[Wright, 1928-1929]

Wright WD, A re-determination of the trichromatic coefficients of the spectral colours, Trans. Opt. Soc. London, 30, pp. 141-164, 1928-1929.

[Wright, 1929-1930]

Wright WD, A re-determination of the mixture curves of the spectrum, Trans. Opt. Soc. London, 31, pp. 201-218, 1929-1930.

[Wright, 1930]

Wright WD, A re-determination of the trichromatic mixture data, Medical Research Council Special Report Series, 139, pp. 5-38, 1930.

[Wyszecki, 2000]

Wyszecki G and Stiles WS, Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd edition, New York, NY: John Wiley, 2000.

[Xue, 2008]

Xue Y, Uniform Color Spaces based on CIECAM02 and IPT Color Difference Equations, Master's Thesis, Rochester, NY: Rochester Institute of Technology, 2008.